



Universidad
Nacional
de Loja



Carrera de Ingeniería en
Sistemas / Computación

Universidad Nacional de Loja

Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables

Carrera de Ingeniería En Sistemas

APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS

**Trabajo de Titulación previa a
la obtención del título de
Ingeniero en Sistemas**

AUTOR:

Roger Alexander Torres Yaguana.

DIRECTOR:

Ing. Francisco Javier Álvarez Pineda. Mg. Sc.

Loja-Ecuador

2022



Loja, 17 de septiembre de 2021.

Ing. Francisco Javier Alvarez Pineda Mg.
DIRECTOR DE TRABAJO DE TITULACIÓN

CERTIFICA:

Que el estudiante señor: **Roger Alexander Torres Yaguana** con C.I: **1103491435**, aprobó la asignatura de trabajo de titulación; respecto al desarrollo del trabajo de titulación denominado **"APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS"**, opción de titulación escogida dentro del periodo académico de culminación de sus estudios superiores en la Carrera de Ingeniería en Sistemas / Computación, siendo las 16h00 del día 30 de septiembre de 2021, se certifica que a la fecha ha cumplido con el cien por ciento (100%) del trabajo de titulación y está en condiciones de continuar con los procesos que correspondan.



**FRANCISCO
JAVIER ALVAREZ
PINEDA**

Francisco Javier Alvarez Pineda.
DIRECTOR DE TRABAJO DE TITULACIÓN

Autoría

Yo, **ROGER ALEXANDER TORRES YAGUANA**, declaro ser autor del presente Trabajo de Titulación y eximo expresamente a la Universidad Nacional de Loja y a sus representantes jurídicos de posibles reclamos o acciones legales por el contenido del mismo. Adicionalmente acepto y autorizo a la Universidad Nacional de Loja, la publicación de mi trabajo en el Repositorio Institucional-Biblioteca Virtual.

Firma: ...



...

Cédula: 1103491435

Fecha: 02 de agosto de 2022

Correo electrónico: roger.torres@unl.edu.ec

Teléfono: 0961665555

Carta de autorización del Trabajo de Titulación por parte del Autor, para la consulta, reproducción parcial o total y publicación electrónica del texto completo.

Yo, **Roger Alexander Torres Yaguana**, declaro ser el autor del trabajo de titulación denominado: **“APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS”**, como requisito para optar el título de: **INGENIERO EN SISTEMAS**; autorizo al sistema Bibliotecario de la Universidad Nacional de Loja para con fines académicos, muestre la producción intelectual de la Universidad, a través de la visibilidad de su contenido de la siguiente manera en el Repositorio Institucional:

Los usuarios pueden consultar el contenido de este trabajo en el Repositorio Institucional, en las redes de información del país y del exterior con las cuales tenga convenio la Universidad.

La Universidad Nacional de Loja, no se responsabiliza por plagio o copia del trabajo de titulación que realice un tercero.

Para constancia de esta autorización, en la ciudad de Loja, a los dos días del mes de agosto del dos mil veinte y dos.

Firma:



.....

Autor: Roger Alexander Torres Yaguana

Cédula: 1103491435

Dirección: Barrio Clodoveo Jaramillo - (Potosí entre Rosario y Asunción)

Correo Electrónico: roger.torres@unl.edu.ec

Teléfono: 072613378

Celular: 0961665555

DATOS COMPLEMENTARIOS:

Director del trabajo de titulación: Ing. Francisco Javier Álvarez Pineda, Mg. Sc.

Dedicatoria

A mi amada Dirlene, a mi familia y a las personas que me quieren...

alex t

Agradecimiento

Primeramente, gracias a todas las personas que han confiado en mí, otorgando una mención especial a la Universidad Nacional de Loja, por dar la oportunidad de aprender los conocimientos con un nivel muy alto de exigencia, a la Carrera de Sistemas y Computación y a todos sus integrantes por permitirme llegar a una meta muy anhelada que es la formatura.

Agradezco también a mi hermano Francisco, por haber sido partícipe de este trabajo de titulación que seguro servirá a las personas que padezcan este tipo de enfermedades y así contribuir con un poquito a ser parte de la solución y no del problema.

Quiero agradecer a los buenos amigos, de los cuales he aprendido mucho en este trajinar de la vida que a veces se torna obscuro pero que siempre hay alguien para poder ayudarlo a uno cuando más precisa.

Finalmente, quiero agradecer a las personas a las cuales no tuve la aceptación y que tengan presente si alguna vez les cause algún mal, sepan perdonar y ante todo entenderme, amigos míos, muchas gracias y que Dios los bendiga siempre.

Índice de Contenidos

Portada.....	i
Certificación	ii
Autoría	iii
Carta de autorización del Trabajo de Titulación por parte del Autor, para la consulta, reproducción parcial o total y publicación electrónica del texto completo.	iv
Dedicatoria.....	v
Agradecimiento	vi
Índice de Contenidos.....	vii
Índice de tablas.....	x
Índice de figuras	xiii
Índice de Anexos.....	xv
1. Título	1
2. Resumen.....	2
2.1. Abstract	3
3. Introducción.....	4
4. Marco teórico.....	6
4.1. Técnicas de aprendizaje automático para el seguimiento de enfermedades relacionadas con la Enfermedad Cardiovascular	6
4.1.1. Árboles de decisión	7
4.1.2. Máquinas de Soporte Vectorial.....	9
4.1.3. Redes Bayesianas	11
4.1.4. Regresión lineal.....	13
4.1.5. Redes Neuronales.....	16
4.1.6. Minería de datos.....	19
4.2. Perfil lipídico, riesgo cardiovascular, diabetes tipo 2 e hipertensión arterial sistémica... 21	
4.2.1. Enfermedad Cardiovascular	21
4.2.2. Riesgo Cardiovascular.....	22
4.2.3. Perfil Lipídico	24

4.2.4. Epidemiología	24
4.2.5. Prevención y Tratamiento	25
4.2.6. Tratamiento No Farmacológico.....	28
4.2.7. Tratamiento Farmacológico	28
4.2.8. Estudio del Corazón de Framingham	31
4.2.9. Metodología Extreme Program XP	32
5. Metodología	34
5.1. Contexto	34
5.2. Proceso	34
5.3. Recursos	35
5.3.1. Científicos	36
5.3.2. Técnicos.	36
5.4. Participantes.	38
6. Resultados	39
6.1. FASE UNO: Realizar una revisión bibliográfica de trabajos relacionados	39
6.1.1. Planificación de la literatura de Proyectos Relacionados.....	39
6.1.2. Análisis bibliográfico.	57
6.1.3. Reporte de la Revisión Bibliográfica	58
6.2. FASE DOS: Recopilación y limpieza de datos para el perfil lipídico y el riesgo cardiovascular	60
6.2.1. Análisis de herramientas informáticas para el proyecto.....	60
6.2.2. Recopilación de la información de los pacientes	74
6.2.3. Limpieza de los datos de las variables	76
6.2.4. Diseño del modelo con aprendizaje automático.....	78
6.3. FASE TRES: Predicción y Validación del Algoritmo Utilizado	81
6.3.1. Proceso	81
6.3.2. Resultados	83
6.3.3. Pruebas	86
7. Discusión	102

7.1. Revisar información bibliográfica de trabajos relacionados con este tipo de proyecto.	102
7.2. Recopilación y limpieza de datos para el perfil lipídico y el riesgo cardiovascular.	103
7.3. Predicción y validación del algoritmo utilizado.....	108
8. Conclusiones	110
9. Recomendaciones	111
10. Bibliografía	112
11. Anexos	127

Índice de tablas

Tabla 1. Técnicas de Aprendizaje Automático	6
Tabla 2. Investigaciones aplicando Árboles De Decisión	8
Tabla 3. Investigaciones aplicando Máquinas de Soporte Vectorial.....	10
Tabla 4. Investigaciones Aplicando Redes Bayesianas.....	12
Tabla 5. Investigaciones aplicando Regresión Lineal	15
Tabla 6. Investigación aplicando Redes Neuronales.....	19
Tabla 7. Investigaciones Aplicando Minería De Datos.....	20
Tabla 8. Cuadro Comparativo de las Metodologías Ágiles (Fuente propia).....	33
Tabla 9. Selección de Documentos	41
Tabla 10. Requerimientos Funcionales	66
Tabla 11. Requerimientos No Funcionales	67
Tabla 12: Representación de las Vistas del Modelo 4+1	68
Tabla 13: Tabla de Historias De Usuario	73
Tabla 14: Criterios de validación de usuario experto.....	87
Tabla 15: Pruebas Del Perfil Lipídico.....	88
Tabla 16: Interpretación de la prueba de rendimiento.....	91
Tabla 17: Variables Del Conjunto Total De Datos Recopilados.....	91
Tabla 18: Estratificación del Riesgo Cardiovascular basado en M. Piepoli.....	105
Tabla 19: Metas de Factores de Riesgo Cardiovascular basado en M. Piepoli	106
Tabla 20: Responsable del proyecto a desarrollar.....	161
Tabla 21: Director del Trabajo de Titulación.....	161
Tabla 22: Definiciones, Acrónimos y Abreviaturas	162
Tabla 23: Características del Usuario Administrador	163
Tabla 24: Características del Usuario Contribuyente.....	164
Tabla 25: Requerimiento Funcional 1	165
Tabla 26: Requerimiento Funcional 2	165
Tabla 27: Requerimiento Funcional 3	166
Tabla 28: Requerimiento Funcional 4.....	166
Tabla 29: Requerimiento Funcional 5.....	166
Tabla 30: Requerimiento Funcional 6.....	167
Tabla 31: Requerimiento Funcional 7.....	167
Tabla 32: Requerimiento Funcional 8.....	167
Tabla 33: Requerimiento Funcional 9.....	168
Tabla 34: Requerimiento Funcional 10.....	168

Tabla 35: Requerimiento No Funcional 1	168
Tabla 36: Requerimiento No Funcional 2	169
Tabla 37: Requerimiento No Funcional 3	169
Tabla 38: Requerimiento No Funcional 4	169
Tabla 39: Requerimiento No Funcional 5	169
Tabla 40: Requerimiento No Funcional 6	169
Tabla 41: Definiciones, Siglas y Abreviaturas	170
Tabla 42: Vistas del Modelo 4 + 1	173
Tabla 43: Historia de Usuario: Autenticar usuario	184
Tabla 44: Historia de Usuario: Cargar los datos cuantitativos de laboratorio del paciente.....	184
Tabla 45: Historia de Usuario: Limpiar los datos cuantitativos del conjunto total de pacientes	185
Tabla 46: Historia de Usuario: Mostrar los datos normalizados en base a métricas requeridas	185
Tabla 47: Historia de Usuario: Ingresar datos para predecir el riesgo cardiovascular	186
Tabla 48: Historia de Usuario: Mostrar el porcentaje del riesgo cardiovascular del paciente ..	186
Tabla 49: Historia de Usuario: Mostrar gráficos estadísticos de los datos cargados en el sistema	187
Tabla 50: Historia de Usuario: Ingresar datos para calcular el perfil lipídico.....	187
Tabla 51: Historia de Usuario: Mostrar el nivel del perfil lipídico del paciente	188
Tabla 52: Historia de Usuario: Enviar por correo los resultados obtenidos del perfil lipídico .	188
Tabla 53: Resumen de las Historias de Usuario.....	189
Tabla 54: Pantalla de inicio de sesión	191
Tabla 55: Pantalla principal del médico administrador	192
Tabla 56: Pantalla principal del listado de datos	192
Tabla 57: Pantalla de administración de limpieza de datos.....	193
Tabla 58: Pantalla de creación de modelo automático	193
Tabla 59: Pantalla principal de riesgo cardiovascular	194
Tabla 60: Pantalla principal de obtención del riesgo cardiovascular	194
Tabla 61: Pantalla de estadísticas de los pacientes.....	195
Tabla 62: Pantalla de estadísticas de las enfermedades asociadas	195
Tabla 63: Pantalla de inserción de variables para el perfil lipídico.....	196
Tabla 64: Pantalla del nivel del perfil lipídico	196
Tabla 65: Pantalla de envío de resultados del perfil lipídico	197
Tabla 66: Requerimientos Funcionales	199
Tabla 67: Caso de prueba nro. 1	200
Tabla 68: Caso de prueba nro. 2.....	200
Tabla 69: Caso de prueba nro. 3.....	201

Tabla 70: Caso de prueba nro. 4.....	202
Tabla 71: Caso de prueba nro. 5.....	202
Tabla 72: Caso de prueba nro.6.....	203
Tabla 73: Caso de prueba nro. 7.....	203
Tabla 74: Caso de prueba nro. 8.....	204
Tabla 75: Caso de prueba nro. 9.....	205
Tabla 76: Caso de prueba nro. 10.....	205

Índice de figuras

Figura 1. Segmento de un árbol de decisión.....	7
Figura 2. Máquina de soporte vectorial	10
Figura 3. Grafo de una red bayesiana	12
Figura 4. Regresión Lineal	14
Figura 5. Esquema de una Red Neuronal Artificial	17
Figura 6. Variable binaria siendo tratada como numérica _Minería de Datos.....	20
Figura 7. Modelo de regresión lineal múltiple con retrodispersión	63
Figura 8. Modelo de Arquitectura MVT (Fuente Propia)	68
Figura 9: Diagrama de caso de uso (Fuente Propia)	69
Figura 10. Diagrama de Clase	70
Figura 11: Diagrama de despliegue (Fuente Propia).....	71
Figura 12: Diagrama de componentes (Fuente Propia)	72
Figura 13: Diagrama de Actividad: Predicción del riesgo cardiovascular y el cálculo del perfil lipídico (Fuente Propia).....	73
Figura 14. Clase principal ClearData	77
Figura 15. Dataset transformado y limpio.....	78
Figura 16. Clase principal de modelo de regresión lineal múltiple	79
Figura 17. Creación del modelo de regresión lineal múltiple para el RCV	80
Figura 18 Dataset limpio del conjunto de pacientes seleccionados	82
Figura 19. Modelo predictivo para el riesgo cardiovascular	84
Figura 20: Variables para el cálculo del perfil lipídico	85
Figura 21. Variables para el cálculo del perfil lipídico.....	86
Figura 22: Codificación de pruebas unitarias del perfil lipídico	88
Figura 23: Estado de las pruebas unitarias	90
Figura 24: Prueba de rendimiento del sitio web CardioZT (Fuente: Apache JMeter).....	90
Figura 26: Importar librerías para el modelo de regresión lineal múltiple.....	94
Figura 27: Clase de Regresión lineal múltiple	95
Figura 28: Obtención del modelo de regresión lineal.....	95
Figura 29: Función inicio de la regresión.....	95
Figura 30: Definir los datos correspondientes a las etiquetas.....	96
Figura 31: Estructura de la regresión lineal múltiple	96
Figura 32: Función de entrenamiento del modelo	97
Figura 33. Porcentaje de éxito del modelo predictivo del riesgo cardiovascular	98
Figura 34. Matriz de confusión Nro. 1.....	99

Figura 35: Matriz de Confusión Nro.2.....	100
Figura 36. Niveles del Perfil lipídico del paciente.....	101
Figura 37. Riesgo cardiovascular con regresión lineal múltiple	109
Figura 38: Base de datos general	158
Figura 39: Base de datos de pacientes diabéticos e hipertensos.....	158
Figura 40: Base de datos general final.....	159
Figura 41: Diagrama de Casos de Uso.....	163
Figura 42: Diagrama de Casos de Uso (Fuente Propia).....	175
Figura 43: Diagrama de Clases (Fuente Propia).....	176
Figura 44: Diagrama de despliegue (Fuente Propia).....	177
Figura 45: Diagrama de componentes (Fuente: Propia)	178
Figura 46: Diagrama de Actividad: Predicción y cálculo de pacientes con DMT2 y HTA.....	179
Figura 47: Diagrama de Actividad: Limpieza de conjunto total de datos de pacientes	179
Figura 48: Diagrama de Actividad: Creación del modelo predictivo	180
Figura 49: Diagrama de Actividad: Obtener el riesgo cardiovascular_Modelo Predictivo	180
Figura 50: Diagrama de Actividad: Obtener el perfil lipídico_Modelo Matemático.....	181

Índice de Anexos

Anexo 1: Entrevista realizada al especialista de Medicina Interna.	127
Anexo 2: Análisis bibliográfico de trabajos relacionados.	131
Anexo 3: Tablas de variables de inclusión.	158
Anexo 4: Documento de Especificación de Requerimientos.	160
Anexo 5: Definición Siglas y Abreviaturas.	170
Anexo 6: Arquitectura del Sistema.	172
Anexo 7: Historias de Usuario.	182
Anexo 8: Prototipo de Interfaz.	190
Anexo 9: Pruebas de Aceptación.	198
Anexo 10: Acta de Acuerdo de Confidencialidad.	208
Anexo 11. Certificaciones del Hospital Ceibos.	214
Anexo 12: Fotografías de la aprobación de la Aplicación Web.	218
Anexo 13: Manuales del programador y usuario.	220
Anexo 14: Certificación y traducción del resumen.	250

1. Título

**APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU
RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES
DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA
DEL HOSPITAL CEIBOS**

2. Resumen

Dentro de las enfermedades cardiovasculares una de las complicaciones más comunes en la sociedad actual es su alto grado de incidencia de eventos con desenlaces severos y/o letales suscitados en su población de convergencia como la aterosclerosis y la dislipidemia, donde el hecho de conocer el estado del perfil lipídico y el riesgo cardiovascular de los pacientes con diabetes tipo 2 e hipertensión arterial sistémica a través de la información recopilada de sus registros permitiría tener un mejor control de los mismos, enfocado en una medida de prevención que evite la desvalorización de la calidad de vida. Así se hace necesario vincular las herramientas tecnológicas informáticas con la medicina humana, por tanto, la interacción entre estas dos áreas permitirá establecer el tipo de relación existente entre el perfil lipídico y riesgo cardiovascular. Por otro lado, este estudio médico-informático se inició recolectando manualmente los datos de los pacientes desde el Hospital General Norte de Guayaquil “Ceibos”, luego realizar una limpieza automática de los mismos, posteriormente se implementó el algoritmo de aprendizaje automático de la regresión lineal múltiple para finalizar con la agrupación de estos tres elementos previos en una aplicación web, donde se obtuvo los resultados siguientes: de un total de 5617 registros de pacientes el 55,59% correspondiente al género femenino y un 44,41 % al masculino 2630 son diabéticos y 3052 son hipertensos. A través de una experimentación se obtuvo una precisión de 85,23% de precisión del modelo y con la matriz de confusión un 90% de efectividad tomando una muestra del conjunto total de datos recopilados.

Así, el objetivo del trabajo de titulación fue la construcción de un sitio web con aprendizaje automático para colaborar conjuntamente con el médico internista, otorgando información fundamental para la toma de decisiones en cuanto a tratamientos para el mejoramiento de la calidad de vida de sus pacientes a través de desarrollo web con aprendizaje automático y determinar con ello que tan bueno o no es su perfil lipídico asociado con su riesgo cardiovascular.

Palabras claves: Aprendizaje automático (AA), Colesterol, Riesgo Cardiovascular.

2.1. Abstract

Among the cardiovascular diseases, one of the most common complications in today's society is its high degree of incidence of events with severe and/or lethal outcomes in its convergence population, such as atherosclerosis and dyslipidemia, where the fact of knowing the state of the lipid profile and cardiovascular risk of patients with diabetes type 2 and systemic arterial hypertension through the information compiled from their records would allow better control of them, focused on a preventive measure that avoids the devaluation of the quality of life. Thus, it is necessary to link computer technology tools with human medicine; therefore, the interaction between these two areas will allow to establish the type of relationship between the lipid profile and cardiovascular risk. On the other hand, this medical-computer study began by manually collecting patient data from the Hospital General Norte de Guayaquil "Ceibos", then performing an automatic cleaning of the same. Later the algorithm of automatic learning of the multiple linear regression was implemented to finish with the grouping of these three previous elements in a web application, where the following results were obtained: of a total of 5617 patient records, 55.59% corresponding to the female gender, and 44.41% to the male gender 2630 are diabetics; and 3052 are hypertensive. Through experimentation, an accuracy of 85.23% of the model was obtained and with the confusion matrix, a 90% effectiveness was obtained, taking a sample of the total set of data collected.

Thus, the objective of the degree work was the construction of a website with automatic learning to collaborate together with the internist doctor, providing essential information for decision-making regarding treatments to improve the quality of life of their patients through web development with automatic learning and thereby determine how good or not their lipid profile is associated with their cardiovascular risk.

Keywords: Machine Learning (ML), Cholesterol, Cardiovascular Risk.

3. Introducción

Está por demás demostrado que la enfermedad cardiovascular inicia a edades muy tempranas, teniendo a la genética como principal protagonista, que silenciosamente va entorpeciendo la arquitectura y función endotelial hasta cuando el daño se hace lo bastante grave como para producir morbimortalidad cardiovascular.

La morbimortalidad por enfermedades cardiovasculares (ECVs) es desde hace varias décadas un problema importante de salud pública y aumenta cada día más en la población económicamente activa. Mientras que las ECVs continúan siendo la causa principal de muerte en los países más desarrollados, la mortalidad por infarto agudo de miocardio parece haber disminuido hasta en un 50% en los años 1990 y 2000. Entre 49 países de Europa y el norte de Asia, más de 4 millones personas mueren cada año por dichas patologías [1].

En América latina, la tasa de mortalidad para las enfermedades cardiovasculares fue de 438/100 000 habitantes; y, en África, Asia Sudoriental y el Mediterráneo Oriental de 650/100 000 habitantes.

En el Ecuador, las enfermedades cardiovasculares forman parte de las diez primeras causas de mortalidad; con una tasa de mortalidad para ambos sexos de 177.6/100 000 habitantes en el año 2000 y de 149.3/100 000 habitantes en el año 2012, siendo en este año para el género masculino de 170.3/100 000 habitantes y para el género femenino de 130.3/100 000 habitantes [2].

Por la importancia que tiene el control de los factores de riesgo cardiovascular, es indispensable conocer el perfil lipídico en la población para poder actuar a través de medidas de prevención y tratar a aquellos pacientes que lo ameriten. El estudio pretende aportar información de los mismos con las patologías más relevantes en nuestro medio como lo son la Diabetes Mellitus Tipo 2 (DMT2) e Hipertensión Arterial Sistémica (HTA) en el Hospital General del Norte de Guayaquil Los Ceibos, vinculando la parte informática (tesista) y la medicina (médico especialista en medicina interna), a través de un estudio retrospectivo, sobre cuál es el perfil lipídico de dichos pacientes y su relación con el riesgo cardiovascular, utilizando herramientas informáticas dirigidas al control y

seguimiento de estas patologías, empleando el aprendizaje automático, a través de uno de los algoritmos supervisados, donde se planeó llevar a cabo la experimentación y consecución de la información idónea y precisa para el desarrollo de la solución informática.

Para finalizar la memoria del trabajo de titulación está estructurado por la sección de revisión de literatura referente al tema abordado para el desarrollo del proyecto, en el área médica, seguida de la sección de materiales y métodos en la cual se detallan las técnicas y materiales utilizadas para la consecución del proyecto, posteriormente la sección de Resultados se vislumbra todo el proceso para la creación del sistema, con la aplicación de la metodología de Programación Extrema. En la sección de Discusión se detalla comparativas de trabajos similares al presente trabajo investigativo con los resultados obtenidos en dichos estudios. Finalmente, en la sección de Conclusiones y recomendaciones las valoraciones ponderables de haber desarrollado el presente proyecto médico- informático de investigación.

4. Marco teórico

4.1. Técnicas de aprendizaje automático para el seguimiento de enfermedades relacionadas con la Enfermedad Cardiovascular

El aprendizaje automático AA tiene diferentes técnicas que ejecutan tareas que habitualmente requieren inteligencia humana, como la percepción visual, el reconocimiento de voz, la toma de decisiones y la traducción entre idiomas. Se necesita que el AA sea regido por principios básicos de ética para sus diferentes aplicaciones en la sociedad [1].

Por ello y en base a la investigación realizada en fuentes bibliográficas se describen las técnicas de aprendizaje automático que se podría implementar para la temática médica investigada (Ver Tabla 1.) con la información de los pacientes con diabetes mellitus tipo 2 e hipertensión arterial sistémica, la misma que servirá para la obtención del modelo predictivo.

Tabla 1. Técnicas de Aprendizaje Automático

Técnicas de Aprendizaje Automático	Referencias Casos de Éxito
1. Árboles de decisión	[3], [4]
2. Máquinas de Soporte Vectorial	[5], [6]
3. Redes Bayesianas	[7], [8]
4. Regresión Lineal	[9], [10], [11]
5. Redes Neuronales Artificiales	[12]
6. Minería de datos	[13], [14]

A continuación, se describen cada una las técnicas mencionadas anteriormente. Para ello cada técnica de Aprendizaje Automático consta de dos secciones, como es Revisión Literaria. La primera sección describe el funcionamiento de la técnica inteligente y la

segunda sección dará a conocer las investigaciones que se han realizado aplicando la técnica en estudio.

4.1.1. Árboles de decisión

4.1.1.1. Revisión Literaria

Los árboles de decisión, proporcionan un conjunto de reglas en el aprendizaje automático para una adecuada toma de medidas para hacer más eficaz en el análisis de los datos y llegar a una correcta solución [15].

Según la metodología empleada, se obtendrán los mejores resultados en la extracción de la información [16]. (Ver Figura 1.)

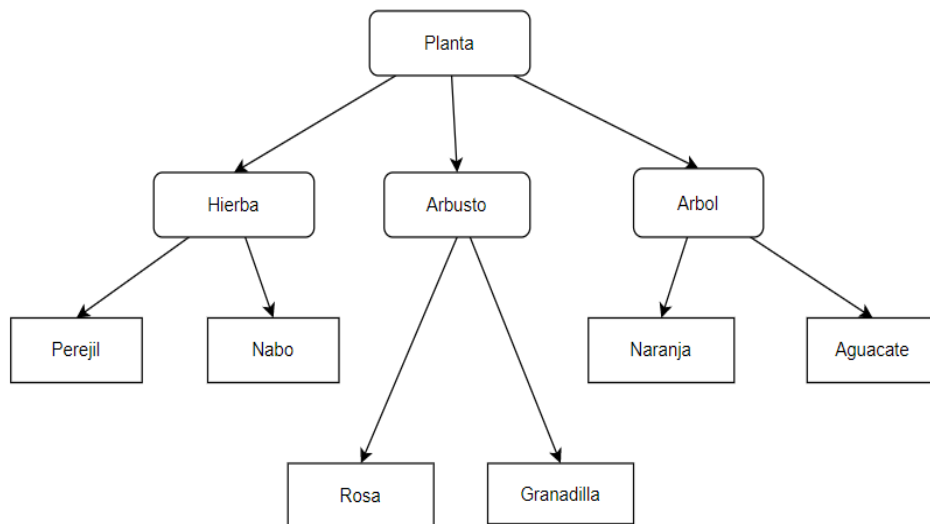


Figura 1. Segmento de un árbol de decisión

Elementos de los Árboles de decisión

Un árbol de decisión está formado por un conjunto de nodos de decisión (interiores) y de nodos-respuesta (hojas):

- ✓ Un nodo de decisión está asociado a uno de los atributos y tiene dos o más ramas que salen de él, cada una de ellas representando los posibles valores que puede tomar el atributo asociado. De alguna forma, un nodo de decisión es como una

pregunta que se le hace al ejemplo analizado, y dependiendo de la respuesta que dé, el flujo tomará una de las ramas salientes.

- ✓ Un nodo respuesta está asociado a la clasificación que se quiere proporcionar, y devuelve la decisión del árbol con respecto al ejemplo de entrada [17].

Casos de Éxito

Existen investigaciones en diversas áreas que hacen uso de los árboles de decisión, una de ellas es el área de la educación (Ver Tabla 2.) el cual se apoyó en esta disciplina para poder explotar la información que existe en el ámbito educativo.

Tabla 2. Investigaciones aplicando Árboles De Decisión

Casos de Éxito	Descripción
1. Predicción de diabetes usando algoritmos de clasificación	Se diseñó un sistema para la detección temprana de la diabetes, donde se evaluaron tres algoritmos de clasificación de aprendizaje automático: árboles de decisión, máquinas de soporte vectorial, naive bayes. Las prestaciones de todos los tres algoritmos se evaluaron en varias medidas como exactitud de precisión, repositorio de aprendizaje, donde los resultados mostraron una precisión alcanzada del 76.3% usando el algoritmo de clasificación de naive bayes [3].
2. Árbol de clasificación para la evaluación del estilo de vida sedentario entre hipertensos	Se basa en la evaluación de estilo de vida sedentaria en personas hipertensas, desarrollando un árbol de clasificación de indicadores clínicos para la correcta predicción del diagnóstico de enfermería “Estilo de vida sedentario” (EVS) en personas con presión arterial alta. Los métodos fueron: un estudio transversal realizado en un centro de atención ambulatoria especializado en hipertensión arterial (HTA) y diabetes de Mellitus en Brasil. La muestra consistió en 285

	<p>personas entre 19 y 59 años de edad diagnosticadas con HTA. El árbol se generó utilizando el algoritmo CHAID (Detección automática de interacción Chi cuadrado), y los resultados fueron la construcción del árbol de decisión que permitió establecer las interacciones entre indicadores clínicos que facilitaron un análisis probabilístico de múltiples situaciones que permitieron cuantificar la probabilidad de que un individuo presente un estilo de vida sedentario. La capacidad predictiva del árbol se estableció en 69.5%. Así, los árboles de decisión ayudan a las enfermeras que cuidan a las personas con HTA en la toma de decisiones al evaluar las características que aumentan la probabilidad de diagnóstico de enfermería EVS, optimizando el tiempo para el diagnóstico [4].</p>
--	--

4.1.2. Máquinas de Soporte Vectorial

4.1.2.1. Revisión Literaria

Con el crecimiento explosivo en datos estructurados y no estructurados, las organizaciones buscan formas de innovar a través del análisis y la ciencia de datos, la disponibilidad de Big Data permitiendo a las organizaciones aprovechar al máximo el análisis de los mismos. Este es el objetivo de las máquinas de soporte vectorial [18].

Los grandes volúmenes de información generados a través de sistemas informáticos pueden representar un gran conocimiento por aprovechar en la creación de nuevos métodos que ayudan a procesar esta información para áreas como la medicina, la industria, la manufactura entre otras. A estos datos se lo emplea para la toma de decisiones, es importante analizarla para hacer un buen manejo de la misma [19]. (Ver Figura 2. , [20])

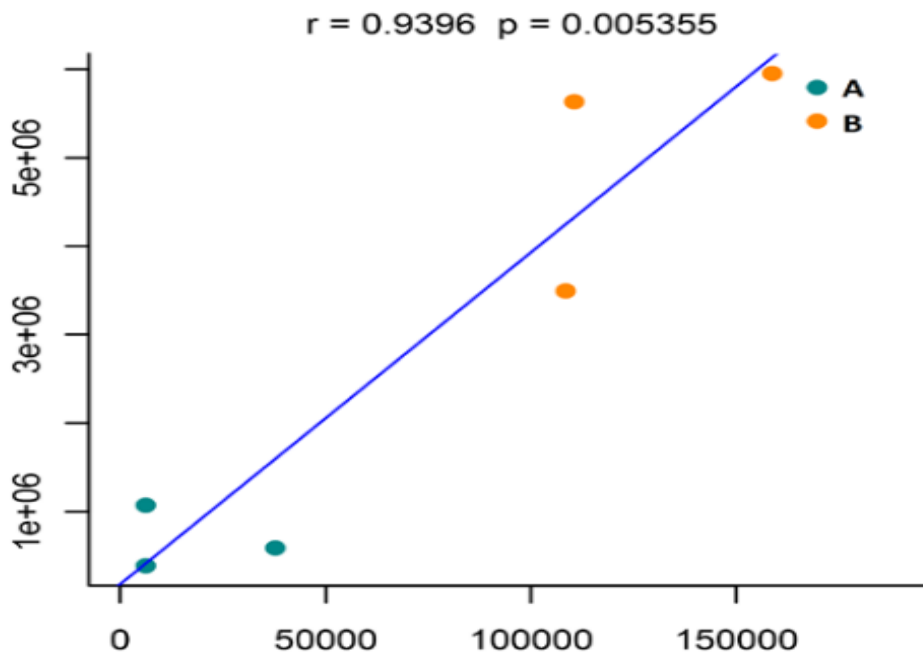


Figura 2. Máquina de soporte vectorial

Casos de Éxito

Se presentan a continuación las investigaciones referentes al procesamiento de la información, acorde con cada estilo de aprendizaje, utilizando la técnica de aprendizaje de la máquina de soporte vectorial (Ver Tabla 3.)

Tabla 3. Investigaciones aplicando Máquinas de Soporte Vectorial

Casos de Éxito	Descripción
1. Un estudio comparativo entre la red neuronal artificial y la máquina de vectores de soporte para el pronóstico del síndrome coronario agudo	Se comparó la capacidad de los modelos de Red Neural Artificial y Máquina de Vectores de Soporte (SVM) en pacientes hospitalizados para clasificar el grado de riesgo de muerte en alto o bajo. Las variables de entrada se seleccionan utilizando el enfoque de envoltura asociado con un filtro de información mutua y dos nuevos filtros propuestos basados en la distancia euclidiana. Los resultados indicaron que el modelo SVM tiene un mejor rendimiento, sus variables relevantes fueron: edad,

	<p>revascularización previa y creatinina, con independencia de los algoritmos de clasificación y los filtros utilizados. Pese a ello los filtros euclidianos además identificaron un segundo grupo importante de variables como lo son: edad, creatinina e hipertensión arterial [5].</p>
<p>2. Implementación de algoritmos de inteligencia artificial para la identificación de pacientes diabéticos utilizando los niveles de lípidos en sangre</p>	<p>La diabetes ocupa el segundo lugar en muertes en México. En este estudio se va a clasificar a estos pacientes utilizando algoritmos de inteligencia artificial: regresión logística, árboles de decisión y máquinas de soporte vectorial, contando con 1019 registros del Centro Médico Siglo XXI, correspondientes a: niveles de colesterol total, lipoproteína de alta densidad, lipoproteína de baja densidad y triglicéridos. La metodología se basó en dos etapas, una de entrenamiento y otra de validación. Donde los resultados evidenciaron un valor de 0.613 hasta 0.727 de AUC, para la detección de pacientes diabéticos [6].</p>

4.1.3. Redes Bayesianas

4.1.3.1. Revisión Literaria

Las redes bayesianas se constituyen en herramientas estadísticas en el campo del aprendizaje automático para afrontar situaciones de investigación con estas características. Una red bayesiana es un modelo gráfico que permite representar las relaciones de dependencia entre un conjunto de variables, mediante el cual se puede dar información muy relevante y ver como se relacionan las variables del dominio, interpretadas en ocasiones como relaciones de causa y efecto [21]. (Ver Figura 3.)

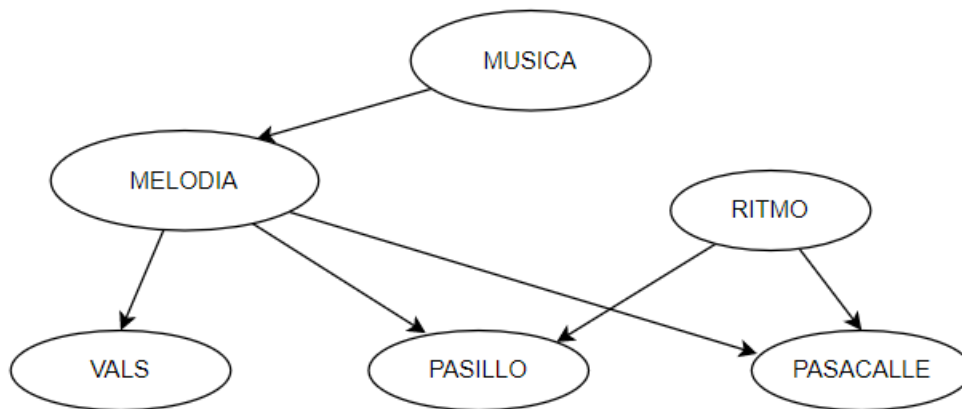


Figura 3. Grafo de una red bayesiana

Casos de Éxito

Se describen diversas investigaciones que han utilizado redes bayesianas para diseñar modelos predictivos que contribuyan a la optimización de pronósticos médicos de la diabetes (Ver Tabla 4.)

Tabla 4. Investigaciones Aplicando Redes Bayesianas

Casos de Éxito	Descripción
1. Sistema Bayesiano para la Predicción de la Diabetes	Trata de un sistema de clasificación Bayesiano para la identificación temprana de la Diabetes Mellitus (DM) analizando variables tales como: número de embarazos, presión arterial diastólica, espesor cutáneo del tríceps, índice de masa corporal, herencia y edad. La metodología propuesta definió y entrenó el sistema propuesto con base en muestras tomadas de pacientes diabéticos y no diabéticos. El sistema se validó en pacientes al azar con proporcionalidad entre los dos grupos; y el grado de aciertos y desaciertos se comparó con un test probado. Los resultados indicaron que en el 87,69% de los casos, el clasificador bayesiano logró detectar

	correctamente esta enfermedad con base en las variables antes mencionadas y cuando se añadió la variable insulina en suero, la detección aumentó hasta el 98,46% [7].
2. Factores sociodemográficos y psicológicos asociados al autocuidado y la calidad de vida en adultos mexicanos con Diabetes Mellitus tipo 2	Buscó identificar las variables sociodemográficas y psicológicas relacionadas con el autocuidado y la calidad de vida en adultos con DMT2. Se utilizó un diseño transversal en una muestra de 60 personas (93 % mujeres) entre 36 y 66 años de edad (M = 54.3, DE = 4.71) adscritas al sistema de salud pública en San Luis Potosí, México. Se midieron las siguientes variables: autocuidado, autoeficacia, conocimiento de la enfermedad, ansiedad, depresión y calidad de vida con las escalas: EECAG, EAG, DKQ-24, AMAS, BDI-II y SF-36. Se utilizó un análisis de regresión lineal para evaluar el impacto de las variables en cuestión. El modelo explicó el 33.9% de la variación del autocuidado a través de las variables depresión (B=-0.27), y autoeficacia (B=0.74). El 56% de variación en Cali de vida relacionada con la salud física se explicó a partir de las variables depresión y autocuidado [8].

4.1.4. Regresión lineal.

4.1.4.1. Revisión Literaria

La regresión lineal sirve para determinar el grado de asociación entre una variable independiente y una dependiente. Es una técnica estadística que analiza la relación entre dos variables cuantitativas que tiene una distribución normal, donde se verifica que existe una relación de forma lineal. Y a medida que la variable independiente está presente, también lo estará en la misma proporción la variable dependiente [22]. (Ver Figura 4.Regresión , [23]).

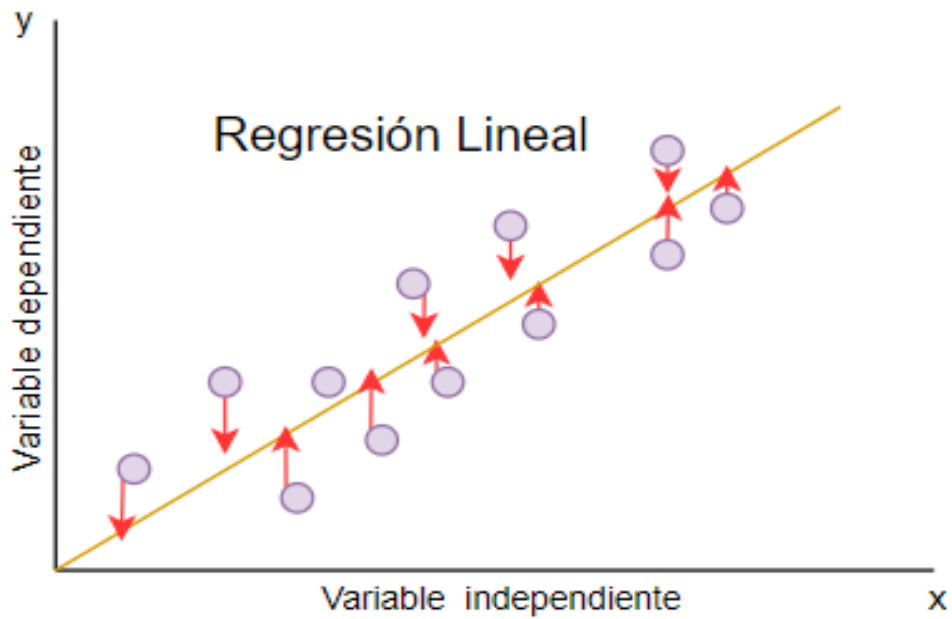


Figura 4. Regresión Lineal

Algunos estudios también pueden clasificarse por el tipo de métodos de correlación empleados, estos pueden ser métodos lineales, como es el caso de la regresión lineal o regresión lineal múltiple, y regresión no lineal [24].

Casos de Éxito

Las siguientes investigaciones describen como se implementa la regresión lineal como apoyo para el análisis y experimentación del desarrollo de enfermedades complejas, haciendo uso de esta técnica estadística (Ver Tabla 5.)

Tabla 5. Investigaciones aplicando Regresión Lineal

Casos de Éxito	Descripción
<p>1. Desarrollo de un algoritmo para determinar el riesgo de muerte en pacientes dentro de una Unidad de Cuidado Intensivo utilizando Regresión Múltiple no Lineal</p>	<p>Los científicos han estudiado el estado de salud de los pacientes, para determinar la severidad de sus enfermedades. En este caso, un algoritmo para la determinación del riesgo de muerte aplicando regresión múltiple no lineal, en comparación con los resultados de los sistemas de clasificación tradicional. Así, de 3479 pacientes estudiados, 420 morían y 3059 sobrevivían, a través de 123 atributos, con la técnica PCA (Análisis de Componentes Principales), utilizando MATLAB [9].</p>
<p>2. Alfabetización en salud y Control de la Diabetes en pacientes de un Hospital Universitario de Argentina.</p>	<p>La AS (Alfabetización en salud) ayuda al seguimiento de los pacientes con enfermedades crónicas, como es la diabetes.</p> <p>Se empleó una regresión lineal para analizar la relación entre la variable dependiente hemoglobina glicosilada (HbA1c) y las predictoras. Las variables que mostraron asociación con $p < 0.1$ en los modelos univariados fueron incorporadas a un modelo de regresión lineal múltiple. Se encontró una inadecuada AS en 94 pacientes (60.3%). La HbA1c fue 8.15%, mientras que los pacientes con adecuada AS, fue 7.15% [10].</p>

<p>3. Control metabólico y rendimiento escolar en niños con diabetes tipo 1</p>	<p>El objetivo de este trabajo fue evaluar la relación entre rendimiento escolar y control metabólico en niños con DM1 (N-DM1) comparando sus resultados con niños de la población general (N-PG). Pacientes y Métodos: Se revisaron datos clínicos de 66 N-DM1. Se compararon las calificaciones de N-DM1 según hemoglobina glicosilada (HbA1c) $< 7,5\%$ y $\geq 7,5\%$ con N-PG del mismo nivel, comuna, tipo de colegio y localidad. Para la comparación de los grupos se utilizó la regresión lineal simple y el test de suma de los rangos de Wilcoxon (Mann y Whitney) previa comprobación de incumplimiento de normalidad con el test de Shapiro-Wilk según el caso. Se consideró un valor de $p < 0,05$ como estadísticamente significativo con una confiabilidad del 95%. Los resultados determinaron que la edad fue $13,4 \pm 2,9$ años, tiempo de evolución DM1 $5,3 \pm 3,2$ años, HbA1c $8,6 \pm 1,9\%$ y controles de glicemia capilar $3,2 \pm 1,2$ veces por día [11].</p>
---	---

4.1.5. Redes Neuronales.

4.1.5.1. Revisión Literaria.

Las redes neuronales artificiales actualmente se han convertido en una rama de estudio multidisciplinaria de las ciencias en general. Inclusive desde antes de que se construyera la primera computadora en 1.946, el hombre se había interesado en el estudio de las redes neuronales y la posibilidad de llevar a la realidad este modelo de forma artificial para diferentes aplicaciones [25].

4.1.5.2. Elementos de una red Neuronal Artificial.

Los elementos básicos de una red neuronal son: entradas, pesos, salidas.

- Las entradas: son las capas que reciben la información desde el exterior.
- Los pesos: son coeficientes que se adaptan dentro de la red para determinar la intensidad de la señal de entrada registrada por la neurona artificial. Además, representan la medida de la fuerza de una conexión de entrada, donde estas fuerzas pueden ser modificadas en respuesta de los ejemplos de entrenamiento de acuerdo a la topología específica o debido a las reglas de entrenamiento.
- Las salidas: donde cada elemento de procesamiento tiene permitido una única salida que puede estar asociada con un número elevado de otras neuronas. Normalmente, la salida es directamente equivalente al valor resultante de la función de activación.

Ciertas topologías de redes neuronales modifican el valor de la función de transferencia para incorporar un factor de competitividad entre neuronas que sean vecinas. Las neuronas tienen permitidas competir entre ellas, inhibiendo a otras neuronas a menos que tengan una gran fortaleza [25], [26]. (Ver Figura 5. Esquema

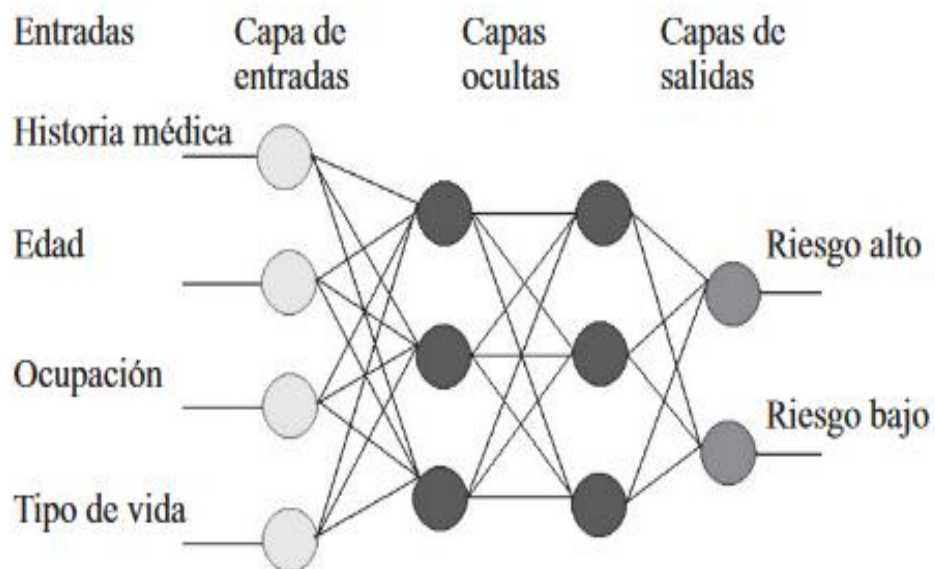


Figura 5. Esquema de una Red Neuronal Artificial

Tipos de aprendizaje básicos de las redes neuronales.

Las redes neuronales para poder aprender se sirven de un algoritmo de aprendizaje, formado por un conjunto de reglas que permiten a la red neuronal aprender a partir de los datos que se le suministran, mediante la modificación de los pesos sinápticos de las conexiones entre las neuronas. Los tipos de aprendizaje de las redes neuronales pueden dividirse básicamente en tres [27]:

- ✓ **Aprendizaje supervisado:** En este tipo de aprendizaje se introducen unos valores de entrada a la red, y los valores de salida generados por esta, donde se comparan con los valores de salida correctos. Si hay diferencias, se ajusta la red en consecuencia.
- ✓ **Aprendizaje de refuerzo:** Para el aprendizaje de refuerzo se ingresan valores de entrada y lo único que se le indica a la red es si las salidas que ha generado son correctas o incorrectas.
- ✓ **Aprendizaje no supervisado:** Respecto al aprendizaje mencionado no existe ningún tipo de guía. Por ende, lo único que puede hacer la red es reconocer patrones en los datos de entrada y crear categorías a partir de estos patrones. Así cuando se le añade algún dato después del entrenamiento, la red será capaz de clasificarlo e indicar en qué categoría lo ha hecho.

Casos de Éxito.

Entre los diversos trabajos con las redes neuronales enfocadas a diversos campos del conocimiento como es el caso de la medicina humana que le dan gran relevancia a esta técnica de aprendizaje automático, la siguiente investigación deja entrever en la tabla que se visualiza a continuación, describiendo el uso de la misma para esta patología como es la retinopatía diabética, dejando ver el poder de predicción para obtener los resultados más idóneos para el control y tratamiento de los pacientes.

Tabla 6. Investigación aplicando Redes Neuronales

Casos de Éxito	Descripción
<p>1. Desarrollo de una aplicación con redes neuronales artificiales para el diagnóstico preventivo de la retinopatía diabética.</p>	<p>Una de las enfermedades más comunes en el Ecuador es la retinopatía diabética, como consecuencia de la diabetes, altos índices de colesterol, hipertensión arterial sistémica. En este trabajo se ha recopilado información necesaria de imágenes de la retina de los pacientes, en el Hospital de Especialidades de las Fuerzas Armadas Nro. 1, para con ello desarrollar un modelo de predicción de retinopatía diabética con el uso de las redes neuronales artificiales. Se realizó un preprocesamiento de imágenes para datos irrelevantes, y en el diseño del modelo se obtuvo un alto índice de precisión de predicciones, valioso aporte para el personal clínico [12].</p>

4.1.6. Minería de datos.

4.1.6.1. Revisión literaria

La minería de datos emerge de los repositorios de datos (Data Warehouse) y de las grandes bases de datos (big Data), fundamentada su extracción de información en la matemática y la estadística. Haciendo uso de la exploración de los datos, la clasificación de datos, la predicción de valores en función de los datos, el modelamiento de las dependencias para resolver problemas de clasificación, el descubrimiento de nuevas reglas y la visualización de sus resultados [15].

4.1.6.2. Métodos de Minería de Datos

Las técnicas más representativas son: las redes neuronales, la regresión lineal, los árboles de decisión, los métodos estadísticos, el agrupamiento o clustering, las reglas de asociación. Según el análisis de los datos, los algoritmos utilizados se clasifican en supervisados o predictivos y no supervisados o del descubrimiento del conocimiento,

donde los primeros predicen un dato desconocido a priori, a partir de otros conocidos y los algoritmos no supervisados descubren patrones y tendencias en los datos [15], [28].

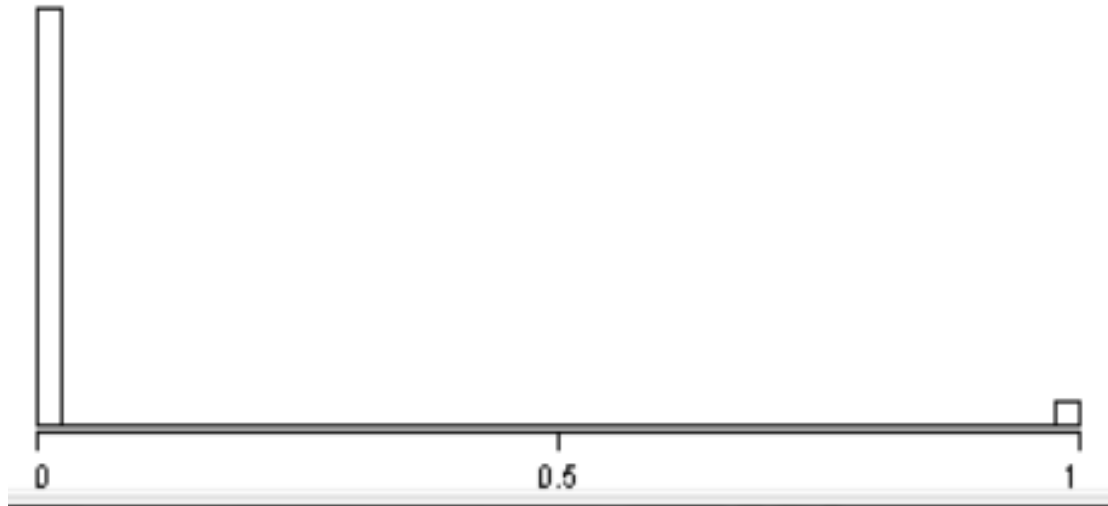


Figura 6. Variable binaria siendo tratada como numérica _Minería de Datos.

Casos de éxito

Existen investigaciones enfocadas a diversos campos que le dan gran relevancia al uso de minería de datos en temas de salud, entre algunas de ellas se tiene las siguientes investigaciones: (Ver

Tabla 7. Investigaciones Aplicando Minería De Datos).

Tabla 7. Investigaciones Aplicando Minería De Datos

Casos de Éxito	Descripción
1. Modelo de simulación y minería de datos para identificar y predecir cambios presupuestales en la atención de pacientes con hipertensión arterial	La problemática de la salud se la ha enfrentado por parte de los países, de acuerdo a su crecimiento económico. Este es un modelo de simulación para determinar el costo que representa para el sistema de seguridad social referente al diagnóstico de la hipertensión arterial.

	<p>Los datos utilizados fueron registros de salud de pacientes, donde se utilizó minería de datos con un modelo de almacenamiento matricial en Matlab para determinar el costo económico de los diagnósticos. Se llegó a determinar que la población diagnosticada representó un 5,7% y el sobrecosto un 43,2%, que correspondieron a valores para poder establecer indicadores posteriores del impacto presupuestal de los registros individuales de salud [13].</p>
<p>2. Funcionalidades de la minería de datos</p>	<p>En este estudio se abordó la metodología y los algoritmos para hacer predicción o clúster de datos de una información específica. La minería de datos tiene su origen en las bases de datos, repositorio de datos y grandes bases de datos, para la extracción de información sustentada en las matemáticas y la estadística. Todo depende de la información obtenida para establecer el modelo para resolver el problema planteado. Así algunas de las aplicaciones de la minería de datos están presentes en la educación, la medicina, la agricultura y otras áreas, donde se utilizará algoritmos de regresión lineal, redes bayesianas, árboles de decisión, lógica e inferencia difusa y redes neuronales [14].</p>

4.2. Perfil lipídico, riesgo cardiovascular, diabetes tipo 2 e hipertensión arterial sistémica

4.2.1. Enfermedad Cardiovascular

El patólogo alemán Marchand, en 1904 introdujo el término aterosclerosis: Athere del griego que significa “papilla”, y Sklero que quiere decir duro. El término hace alusión a

los componentes blando y duro de la placa aterosclerótica que identifica a esta entidad. Como una categoría diagnóstica, enfermedad cardiovascular (ECV) incluye cuatro áreas principales: la enfermedad coronaria, la enfermedad cerebrovascular, la enfermedad arterial periférica, la aterosclerosis aórtica y torácica o aneurisma de aorta abdominal [29].

4.2.2. Riesgo Cardiovascular

Las guías de práctica clínica recomiendan como paso inicial dentro de la valoración de los pacientes evaluar el riesgo cardiovascular absoluto y relativo, ante los hallazgos, cuanto mayor sea su riesgo, se debe intensificar tanto la promoción de los cambios en el estilo de vida y el tratamiento farmacológico, con el objetivo de reducir las probabilidades de un desenlace fatal a corto, mediano o largo plazo.

Para ello existen variedad de sistemas de evaluación del riesgo, uno de ellos es el sistema SCORE (Systematic Coronary Risk Estimation) por sus siglas en inglés, que puede ser reajustado para utilizarlo en diferentes poblaciones no europeas.

En términos generales pacientes con enfermedad aterosclerótica (ASCVD) documentada, Diabetes Mellitus tipo 1 (DM1), Diabetes Mellitus tipo 2 (DM2), enfermedad renal crónica (ERC) son condiciones de muy alto o alto riesgo cardiovascular. Por otro lado, en personas aparentemente sanas, el uso de un sistema de estimación de riesgo estimado a 10 años para el desarrollo de un primer evento aterosclerótico fatal puede revelar situaciones que en conjunto podrían evidenciar altos niveles de riesgo cardiovascular.

Un problema particular se relaciona en los jóvenes con altos niveles de factores de riesgo; un riesgo absoluto bajo puede ocultar un riesgo relativo muy alto que requiere al menos un asesoramiento intensivo sobre el estilo de vida. Para motivar a los jóvenes menores de 40 años a no retrasar el cambio de su estilo de vida poco saludable, puede ser útil una estimación de su riesgo relativo, que ilustra que los cambios pueden reducir sustancialmente dicho riesgo.

A continuación, se muestra una categorización de riesgo cardiovascular en 4 niveles, publicado en las guías europeas del año 2019:

a) Muy Alto Riesgo Cardiovascular

En personas con alguno de los siguientes factores:

- ✓ Enfermedad aterosclerótica documentada por clínica y/o imágenes, esto incluye síndrome coronario agudo (infarto agudo de miocardio o angina inestable), angina estable, revascularización coronaria, stroke e infarto cerebral transitorio, enfermedad arterial periférica.
- ✓ Diabetes Mellitus DM con daño de órgano blanco, o al menos tres factores de riesgo mayor, o inicio temprano de Diabetes Mellitus tipo 1, o duración mayor de 20 años de la enfermedad.
- ✓ Enfermedad renal crónica severa (índice de filtrado glomerular estimado menor de 30 ml/min/1.73 m²).
- ✓ Un SCORE calculado mayor o igual a 10% estimado a 10 años para riesgo de enfermedad cardiovascular fatal.
- ✓ Insuficiencia cardíaca con ASCVD u otro factor de riesgo mayor.

b) Alto Riesgo Cardiovascular

En personas con los factores:

- ✓ Factores de riesgo individual marcadamente elevados, en particular colesterol total (COL-T) mayor a 310 mg/dl, LDL-C mayor a 190 mg/dl, o presión arterial mayor o igual a 180/110 mmHg.
- ✓ Insuficiencia cardíaca sin otros factores de riesgo mayor.
- ✓ DM sin evidencia de daño de órgano blanco, con DM mayor o igual a 10 años desde el inicio de enfermedad y otro factor de riesgo adicional.
- ✓ Enfermedad renal crónica moderada (índice de filtrado glomerular estimado entre 30 - 59 ml/min/1.73 m²).
- ✓ Un SCORE calculado mayor o igual a 5% y menor de 10% estimado a 10 años para riesgo de enfermedad cardiovascular fatal.

c) Moderado Riesgo Cardiovascular

Paciente joven (DM1 menor de 35 años de edad; DM2 menor de 50 años de edad) con DM de duración menor de 10 años de evolución de la enfermedad, sin otros factores de riesgo.

Un SCORE calculado mayor o igual a 1% y menor de 5% estimado a 10 años para riesgo de enfermedad cardiovascular fatal.

d) Bajo Riesgo Cardiovascular

Un SCORE calculado menor de 1% estimado a 10 años para riesgo de enfermedad cardiovascular fatal [30].

4.2.3. Perfil Lipídico

Las dislipidemias son un conjunto de enfermedades asintomáticas, que son originadas por concentraciones anormales de lipoproteínas sanguíneas. El sobrepeso y la obesidad coadyuvan al desarrollo de dislipidemias. Los cambios sociales e individuales que acompañan al proceso de urbanización claramente juegan un rol en el incremento de niveles de colesterol en la población urbanizada. Transición debida principalmente al aumento en el consumo de grasas en la dieta, principalmente de origen animal y aceites vegetales procesados, sumado a una menor actividad física.

No existe un umbral específico en el cual desaparece el riesgo cardiovascular, aunque es bastante menor cuando encontramos un LDL-C bajo y HDL-C elevado, a más de la presencia de subfracciones de lípidos [31].

4.2.4. Epidemiología

Las enfermedades cardiovasculares, de las cuales la enfermedad cardiovascular aterosclerótica es su principal componente, es responsable de más de cuatro millones de muertes en Europa cada año, aproximadamente ocurre en mujeres 2.2 millones más que en varones 1.8 millones; y, sin embargo, la muerte cardiovascular antes de los 65 años de edad es mayor en hombres con 490 000 casos versus mujeres 193 000 casos [32].

El riesgo de por vida de enfermedad cardiovascular en general se acerca al 50% de las personas mayores de 30 años sin enfermedad cardiovascular conocida [33].

La enfermedad coronaria representa aproximadamente desde un tercio a la mitad del total de casos de enfermedades cardiovasculares. El Framingham Heart Study puso de manifiesto el riesgo de enfermedades cardiovasculares en 7733 personas, de entre 40 y 94

años de edad que estaban inicialmente libres de enfermedad. El riesgo de por vida para las personas de 40 años fue 49% en hombres y 32% en mujeres. Incluso aquellos que estaban libres de enfermedad coronaria a los 70 años tenían un riesgo de por vida de 35% en hombres y 24% en mujeres [33], [34].

Junto con la disminución de la mortalidad asociada con el evento inicial por enfermedad cardiovascular, su prevalencia está aumentando rápidamente en algunos países desarrollados [35].

En un estudio de 156.424 personas de 17 países (3 de ingresos altos, 10 de ingresos medios, 4 de menores ingresos), la puntuación de riesgo InterHeart fue más alta en los países de ingresos altos y baja en los de menores ingresos [36].

En EE.UU., las encuestas han mostrado pocos cambios en los niveles de lípidos sin tratamiento en adultos. Sin embargo, entre los años 1988 a 1994 y 2005 a 2010, la prevalencia del tratamiento se ha incrementado y ha dado lugar a mejoras en aquellos con los niveles más altos de colesterol. La proporción de personas con un colesterol total mayor o igual a 240 mg/dl ha disminuido del 20,5% al 14,2%; mientras que, aquellos con niveles de colesterol total entre 200 y 239 mg/dl o tratados para meta han aumentado del 32,8% al 39,9%.

En contraste, los niveles internacionales de colesterol han cambiado muy poco; entre 1980 y 2008, el nivel medio de colesterol total fue de 179 mg/dl en los hombres y 184 mg/dl en las mujeres, con una disminución de 3,1 mg/dl por década [37].

4.2.5. Prevención y Tratamiento

La prevención es definida como un conjunto de acciones coordinadas, dirigidas a la población o a nivel individual, destinadas a eliminar y/o minimizar el impacto de las enfermedades cardiovasculares y las secuelas resultantes. Actualmente cada vez más pacientes están sobreviviendo a un primer evento cardiovascular y se convierten en pacientes de alto riesgo de recurrencias. Adicionalmente, la prevalencia de algunos factores de riesgo como Diabetes Mellitus y Obesidad están aumentando. En general la premisa consiste en promover estilos de vida saludables [38].

Las estrategias de manejo del paciente dislipidémico deben basarse en la evaluación del riesgo cardiovascular global del paciente.

Las estatinas están contraindicadas en las mujeres embarazadas o en las que pueden quedar embarazadas.

El médico tiene que identificar en primera instancia al tipo de paciente que va a tratar; el manejo de los factores de riesgo para la enfermedad cardiovascular aterosclerótica (ECV), de los cuales el colesterol elevado de lipoproteínas de baja densidad (LDL-C) es la base, se llama prevención primaria si este proceso se realiza en alguien que no ha experimentado previamente un evento en un vaso sanguíneo aterosclerótico; los niveles de Colesterol reducen el riesgo en pacientes con y sin ECV.

En otro contexto, los pacientes con enfermedad cardiovascular conocida tienen un alto riesgo de futuros eventos cardiovasculares a cuya terapéutica se la conoce como prevención secundaria.

Se ha demostrado que las estatinas, fibratos, la ezetimiba y los inhibidores de PCSK9 reducen el riesgo de eventos cardiovasculares adversos. Entre estas tres, las estatinas son las mejor estudiadas.

Las metas de control terapéutico del perfil lipídico actualmente tienen pocas variaciones entre las distintas sociedades científicas en el mundo, a grandes rasgos se promueve la idea en la cual entre mayor sea el riesgo cardiovascular de un paciente, menor tiene que ser sus valores de LDL-C.

A continuación, se enuncian las nuevas recomendaciones propuestas por la Sociedad Europea de Cardiología (ESC) en conjunto con la Sociedad de Aterosclerosis Europea (EAS) en las guías emitidas en el 2019:

- ✓ La evaluación de la carga de placa arterial (carotídea y / o femoral) en la ecografía arterial debe considerarse como un modificador del riesgo en los individuos con riesgo bajo o moderado.

- ✓ La evaluación de la puntuación del calcio arterial coronario (CAC) con Tomografía debe considerarse como un modificador del riesgo en la evaluación del riesgo cardiovascular de personas asintomáticas con riesgo bajo o moderado.
- ✓ La medición de lipoproteína a (Lp a) debe considerarse al menos una vez en la vida de cada persona adulta para identificar a aquellos con niveles de Lp (a) heredados muy altos > 180 mg / dL que pueden tener un riesgo de por vida de ASCVD equivalente al riesgo asociado con el de hipercolesterolemia familiar heterocigótica.
- ✓ En pacientes de alto riesgo o más con triglicéridos entre 135-499 mg / dl a pesar del tratamiento con estatinas, ácidos grasos poliinsaturados debe considerarse en combinación con estatinas.
- ✓ En prevención primaria, para las personas con hipercolesterolemia familiar (HF) con riesgo muy alto, una reducción de LDL-C mayor o igual al 50% desde el valor inicial y una meta de LDL-C de <55 mg / dl debe considerarse.
- ✓ Se recomienda el tratamiento con estatinas como prevención primaria, según el nivel de riesgo, en adultos menores de 75 años de edad.
- ✓ Se puede considerar el inicio del tratamiento con estatinas para la prevención primaria en personas mayores de 75 años, si tienen alto riesgo o más.
- ✓ En pacientes con diabetes mellitus tipo 2 (DM2) con riesgo muy alto, se recomienda una reducción de LDL-C de mayor o igual al 50% con respecto al valor inicial y un objetivo de LDL-C <55 mg / dl. En pacientes con DM2 de alto riesgo, se recomienda una reducción de LDL-C menor o igual al 50% desde el valor inicial y un objetivo de LDL-C <70 mg / dl. Las estatinas se recomiendan en pacientes con DM1 que tienen un riesgo alto o muy alto.
- ✓ Se debe considerar la intensificación de la terapia con estatinas antes de la introducción de la terapia combinada. Si no se alcanza el objetivo, se debe considerar la combinación de estatinas con ezetimiba.

- ✓ No se recomienda la terapia con estatinas en pacientes pre menopáusicas con DM que estén considerando quedarse embarazadas o no estén usando un método anticonceptivo adecuado.
- ✓ Para pacientes que presentan un síndrome coronario agudo (SCA) y cuyos niveles de LDL-C no están en el objetivo a pesar de que ya toman una dosis de estatinas máxima tolerada, se debe considerar la posibilidad de ezetimiba, agregar un inhibidor de PCSK9 poco después del evento (si es posible, durante la hospitalización por el evento de SCA) [30].

4.2.6. Tratamiento No Farmacológico

Un estilo de vida saludable es la base de la pirámide en el manejo de pacientes cardiovasculares que básicamente se rigen al aspecto nutricional, actividad física diaria, abandono del alcohol y del cigarrillo. El propósito es que el paciente mantenga un estilo de vida saludable, controlando y/o eliminando los factores de riesgo cardiovascular modificables. Es la estrategia costo/beneficio rentable y que ha demostrado reducir los lípidos séricos en una magnitud similar a la que se obtiene con dosis bajas de una estatina, LDL-C del 10% - 15%, triglicéridos 20% - 30%, con un factor adicional el cual consiste en incremento no despreciable del HDL-C; y secundariamente un control más adecuado de tensión arterial y glucemia [30].

4.2.7. Tratamiento Farmacológico

Debe sostenerse en la evaluación de riesgos que tengan en cuenta más que el nivel básico de LDL-C. El riesgo cardiovascular debe calcularse utilizando modelos/calculadoras de riesgo cardiovascular validados. Los pacientes y sus proveedores pueden decidir si una reducción del riesgo relativo del 20% al 30%, que es una expectativa razonable para la terapia con estatinas, se traduce en una reducción del riesgo absoluta lo suficientemente grande como para justificar el costo/beneficio de la terapia diaria.

Aunque recomendamos el cálculo del riesgo de ECV para la mayoría de los pacientes, esto puede ser innecesario para las personas con un C-LDL > 190 mg / 100 ml, ya que generalmente se prescribe el tratamiento con estatinas. Para las mujeres jóvenes en edad fértil que desean quedar embarazadas, se discute cuidadosamente los posibles beneficios

y riesgos con el paciente; las estatinas están contraindicadas en las mujeres embarazadas o en las que pueden quedar embarazadas.

Se tiene que identificar en primera instancia al tipo de paciente que vamos a tratar; el manejo de los factores de riesgo para la enfermedad cardiovascular aterosclerótica (ECV), de los cuales el colesterol elevado de lipoproteínas de baja densidad (LDL-C) es la base, se llama prevención primaria si este proceso se realiza en alguien que no ha experimentado previamente un evento en un vaso sanguíneo aterosclerótico; los niveles de Colesterol reducen el riesgo en pacientes con y sin ECV.

En otro contexto, los pacientes con enfermedad cardiovascular conocida tienen un alto riesgo de futuros eventos cardiovasculares a cuya terapéutica se la conoce como prevención secundaria. Las intervenciones de prevención secundaria están dirigidas a factores de riesgo modificables conocidos para los eventos de ECV como el tabaquismo, la hipertensión, la diabetes y el LDL-C.

Todos los adultos mayores de 20 años deben tener una evaluación de su riesgo de enfermedad cardiovascular. Ofrecemos terapia con estatinas a la mayoría de los pacientes con un riesgo de enfermedad cardiovascular de 10 años de 7.5% a 10% o más.

Tratamiento inicial: se ha demostrado que las estatinas, la ezetimiba y los inhibidores de PCSK9 reducen el riesgo de eventos cardiovasculares adversos. Entre estas tres, las estatinas son las mejor estudiadas. Son más efectivos que el ezetimibe y significativamente menos costosos y fáciles de usar que los inhibidores de PCSK9. Además, se han utilizado con un excelente perfil de seguridad durante más de 30 años.

En pacientes estables (pacientes ambulatorios), se trata con estatinas de alta intensidad (atorvastatina 40 a 80 mg o rosuvastatina 20 a 40 mg al día). Algunos expertos comienzan con la dosis más baja de estatinas de alta intensidad; en pacientes con síndrome coronario agudo muy reciente, comenzamos con atorvastatina 80 mg o rosuvastatina 20 mg.

La falla de un paciente para lograr la reducción esperada de LDL-C con dosis máximas de atorvastatina o rosuvastatina puede deberse a una mala adherencia al tratamiento y recomendaciones de estilo de vida o una respuesta variable a la terapia.

Tratamiento de segunda línea: la evidencia de varios ensayos sugiere que la adición de un segundo fármaco reductor de LDL-C proporciona una reducción adicional y significativa en el riesgo de eventos CVD.

Hay pacientes que no pueden tomar estatinas de alta intensidad o cuyo LDL-C permanece significativamente por encima de 70 mg/dl con las estatinas de alta intensidad. La disminución adicional de LDL-C con terapia sin estatinas puede estar indicada para estos pacientes. Dos terapias demostraron mejores resultados cardiovasculares cuando se añadieron (individualmente) a la estatina: ezetimiba e inhibidores PCSK9. En muchas situaciones, elegimos ezetimibe antes de PCSK9 como terapia de segunda línea por razones de costo.

La terapia con estatinas más ezetimibe en comparación con la estatina disminuye modestamente el riesgo de eventos CVD, pero no la mortalidad, en pacientes con un síndrome coronario agudo. En el ensayo IMPROVE-IT, los valores finales de LDL-C fueron 53.7 y 69.5 mg/dl (1.4 versus 1.8 mol/l) en los grupos de estatina y ezetimiba y estatinas, respectivamente.

Terapia de tercera línea: para pacientes con CVD establecida y un LDL-C superior a 70 mg /dl con estatinas más ezetimiba, considere agregar un anti PCSK9. El beneficio absoluto será mayor en los pacientes con CVD de mayor riesgo, como aquellos con un síndrome coronario agudo. Si el paciente puede recibir tratamiento con anti PCSK9, el médico puede reevaluar el valor de usar ezetimiba en un régimen de tres medicamentos versus discontinuar ezetimiba y emplear el régimen de dos fármacos de una estatina más anti PCSK9.

Otras terapias posibles para pacientes que no alcanzaron su objetivo de LDL-C utilizando terapias de primera y segunda línea, se debe derivar a un médico especializado en el tratamiento de trastornos complejos de los lípidos. Mientras que otras terapias como

medicamentos como mipomersen o lomitapide o aféresis LDL son potencialmente valiosas, su costo puede evitar que se prescriban [30].

4.2.8. Estudio del Corazón de Framingham

En 1948 se puso en marcha el Framingham Heart Study por parte del Servicio de Salud Pública de EEUU. Con el objetivo de estudiar la epidemiología y los factores de riesgo de la enfermedad cardiovascular [39]. Se eligió la ciudad de Framingham, situada a 32 km al oeste de la ciudad de Boston, estado de Massachusetts, en la costa este de la nación norteamericana, porque en ella ya se había realizado con éxito un estudio de base poblacional sobre la tuberculosis en 1918, y, además por su proximidad a los principales centros médicos de Boston; la presencia de empresas grandes y el apoyo prestado por la comunidad médica y la sociedad civil que estaban bien informadas y se mostraban muy colaboradoras con el proyecto.

Cuatro años después del inicio del Framingham Heart Study, la investigación identificó que el colesterol elevado y la presión arterial alta eran factores importantes en cuanto a la aparición de la enfermedad cardiovascular. En los años siguientes, el estudio de Framingham y otros estudios epidemiológicos contribuyeron a identificar otros factores de riesgo, que ahora se consideran ya clásicos.

Al acuñar la expresión «factor de riesgo», el Framingham Heart Study facilitó un cambio en el ejercicio de la medicina. En la actualidad, se define factor de riesgo como un elemento o una característica mensurable (medible) que tiene una relación causal con un aumento de frecuencia de una enfermedad y es un factor predictivo independiente y significativo del riesgo de contraer una enfermedad.

La primera cohorte la conformaron 5.209 habitantes sanos, cuyas edades oscilaban entre 30 y 60 años, que se incorporaron al estudio en 1948, para la realización de exámenes bianuales que han continuado desde entonces. En 1971, se seleccionó a 5.124 hijos e hijas (y sus cónyuges) de la cohorte inicial, para su inclusión en el Estudio de la descendencia. Finalmente, en 2002, un total de 4.095 participantes se incorporaron a la cohorte de tercera generación del estudio [40].

Entre los factores asociados a la enfermedad cardiovascular se logró identificar a los lípidos, y dentro de ellos al *LDL-C* que en concentraciones elevadas desde etapas tempranas de la vida genera daño en el endotelial vascular como un proceso continuo de deterioro de la salud del individuo [41]; otro factor identificado fruto de este estudio fue la *HTA* con lo cual se logró demostrar que la presión arterial sistólica y la diastólica tienen una asociación continua, independiente, gradual y positiva con los parámetros de evolución cardiovascular [42]; el *tabaquismo* no era aceptado como causal de cardiopatía, El Framingham Study y el Albany Cardiovascular Health Center Study demostraron que los fumadores presentaban un aumento del riesgo de infarto de miocardio o muerte súbita; y, que el riesgo estaba relacionado con el número de cigarrillos consumidos al día, y los ex fumadores tenían una morbimortalidad por Enfermedad Coronaria similar a la de los individuos que nunca habían fumado [43]. La diabetes se asocia a un aumento de 2-3 veces la probabilidad de aparición de una ECV [44], y este aumento es mayor en las mujeres que en los varones [45].

La obesidad es un trastorno metabólico crónico asociado a numerosas comorbilidades, como Enfermedad Coronaria, Enfermedad Cardiovascular, diabetes mellitus tipo 2, hipertensión arterial, determinados cánceres y apnea durante el sueño. Además, la obesidad es un factor independiente del riesgo de mortalidad por todas las causas, relación que fue identificada por los investigadores de Framingham hace 40 años [46].

4.2.9. Metodología Extreme Program XP

La Programación Extrema XP, es una de las más populares en el desarrollo de software, teniendo como constante siempre una interacción entre el equipo de desarrollo y los usuarios o clientes finales, aquí se desarrollan pruebas para cada función a implementarse [47].

Algunas características de la Programación Extrema son [48]:

- ✓ La comunicación entre los miembros del proyecto es uno de los principales fundamentos de XP, para un buen escenario de trabajo.
- ✓ La satisfacción del cliente es una prioridad, otorgando sencillez en la medida de lo posible.

- ✓ La retroalimentación permite la construcción del proyecto de manera conjunta entre el equipo de desarrollo y el cliente.
- ✓ El coraje de todo el equipo de desarrollo y ejecución del proyecto estarán prestos a los cambios de los requerimientos, dando su mayor esfuerzo.

Tabla 8. Cuadro Comparativo de las Metodologías Ágiles (Fuente propia)

METODOLOGÍAS ÁGILES	
Metodología Scrum	Divide el trabajo de desarrollo en etapas pequeñas: análisis, desarrollo y testing.
Metodología Extreme Program _XP	Establece una relación entre el desarrollador y el cliente, todo el tiempo.
Metodología Iconix	Tiene un control exacto del ciclo de vida del proyecto a través de métodos de orientación a objetos.
Metodología Kanban	Consolida el proyecto en tareas: pendientes, en proceso o terminadas; mejorando la productividad.
Metodología Crystal Clear	Es una familia de metodologías con un código genético común, para un grupo de 3 a 8 personas.
Metodología Mobile-D	La prioridad de Mobile-D es la información en tiempo real del producto para el usuario final.

5. Metodología

De acuerdo con el Reglamento de Régimen Académico que rige a las Instituciones de Educación Superior del Ecuador, en el artículo 21, numeral 3, se estipula que un Trabajo de Titulación (TT) se basará en procesos de investigación e intervención. Por otro lado, todo TT deberá consistir en una propuesta innovadora que contenga como mínimo una investigación exploratoria y diagnóstica, además de acuerdo con el artículo 72 del mismo reglamento, a nivel de grado es de carácter exploratorio y descriptivo [49].

Para dar inicio al desarrollo del presente TT se realizó un análisis exploratorio de trabajos relacionados con el objeto de estudio para así tener en claro el problema a solucionar, posteriormente se procedió a definir el contexto donde se desarrolló el TT, seguidamente se definió el proceso en base a las actividades establecidas de cada objetivo específico, luego se definen los recursos que se utilizaron para el desarrollo del TT y finalmente se detallan los participantes que intervinieron en la ejecución del mismo.

Mediante la investigación descriptiva se logró definir la metodología de desarrollo de software y mediante esto lograr cumplir a cabalidad el objetivo general del TT.

5.1. Contexto

La presente investigación se lo realizó en el Hospital General Norte de Guayaquil, en el departamento de Consulta Externa, donde se plantearon los lineamientos necesarios para el desarrollo del TT, conjuntamente con el docente de la materia y el tutor. Así mismo se tuvo el apoyo del personal adyacente al departamento como enfermeras y personal de seguridad del Hospital.

5.2. Proceso

Para dar cumplimiento al objetivo de “Elaborar un sitio web para determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital General Norte de Guayaquil Los Ceibos”, se siguieron una serie de actividades planteadas por cada uno de los objetivos específicos. Seguidamente, se detallan los objetivos específicos con sus respectivas actividades:

1. Analizar información bibliográfica de trabajos relacionados con este tipo de proyecto

a. Mediante la entrevista se pudo recolectar información precisa acerca de la problemática de salud de los pacientes diabéticos e hipertensos de sus enfermedades acaecidas y así poder iniciar al desarrollo del TT, (Ver Anexo1).

b. Con la ayuda de una revisión literaria de trabajos similares al objeto de estudio se pudo obtener y complementar los requerimientos de software, para luego definir el documento de Especificación de Requerimientos de Software (ERS) según la normativa IEEE 830. (Ver Anexo 4).

c. En base a los requerimientos obtenidos en la ERS se procedió a diseñar la arquitectura del software tomando como base el modelo 4+1 de Kruchten. (Ver Anexo 6).

2. Recopilación y limpieza de datos para el perfil lipídico y el riesgo cardiovascular.

a. Definir las variables de estudio con criterios de los expertos en medicina interna de la institución hospitalaria. (Ver Anexo 3).

b. Recopilar la información de los pacientes desde el programa AS400 del Hospital Ceibos en hojas de Excel (Ver Anexo 13).

c. Limpiar los datos recopilados manualmente con funciones automáticas y las librerías pandas, numpy y scikit-learn para el análisis y tratamiento de la data.

3. Predicción y validación del algoritmo utilizado

a. Construir la aplicación con el modelo de aprendizaje automático de la regresión lineal múltiple para el conjunto de datos recopilados.

b. Realizar las pruebas de aceptación del software a usuarios de la aplicación informática. (Ver Anexo 9).

5.3. Recursos

De acuerdo a la pregunta de investigación para su respuesta conforme a los objetivos planteados se utilizaron los siguientes recursos, detallados a continuación.

5.3.1. Científicos

- ✓ **Método Deductivo:** Mediante este método el investigador pudo derivar problemas generales que han sido descritos en trabajos equivalentes a la problemática del trabajo de investigación que hayan sido solucionados a través de una herramienta informática, como consecuente base principal para iniciar el presente trabajo.
- ✓ **Método Inductivo:** Este recurso científico ayudó al investigador a conocer los principales problemas y necesidades que actualmente tiene el personal de salud de la Consulta Externa del Hospital Ceibos, partiendo de conocimientos específicos que posteriormente se generalizaron para todos los médicos.
- ✓ **Entrevista:** A través de ella, el investigador pudo instaurar un diálogo con las partes complementarias del vigente trabajo investigativo, lo cual sirvió para la recopilación de requerimientos e instauración de los lineamientos precisos para iniciarlo (Ver ANEXO 1: Modelo de Entrevista realizada al especialista de Medicina Interna).
- ✓ **Método Analítico:** Técnica que ayudó a ordenar la información recogida, para después analizar cada uno de sus elementos y la relación que existe entre ellos. Lo que ayudo a definir visiblemente los requerimientos de software y sus relaciones con el modelo de dominio (Ver Sección 6. Resultados apartados 3. Elaboración del documento de requerimientos de software).
- ✓ **Observación Directa:** Esta práctica sirvió para recopilar información necesaria mediante la observación, y así comprender mejor el problema en estudio y tener una idea clara para la representación de los requerimientos que el software requiere. (Ver Sección de Resultados Apartado 3. Elaboración del documento de requerimientos)

5.3.2 Técnicos.

5.3.2.1. IEEE 830.

Con esta normativa de la IEEE utilizada para la realización de la Especificación de Requerimientos de Software, que sirvió para delimitar, contrastar y enfocar cada una de las necesidades que tiene los pacientes diabéticos tipo 2 e hipertensos del Hospital Ceibos,

para luego complementarlos con los requerimientos encontrados en trabajos similares y así dar una solución óptima a la problemática del trabajo de titulación.

5.3.2.2. Metodologías para el desarrollo de software

Para el desarrollo del presente trabajo de titulación se utilizó la metodología de desarrollo ágil Extreme Programming (XP). A continuación, se detallan cada una de sus fases correspondientes.

5.3.2.3. Metodología XP

La metodología ágil de la Programación Extrema (XP) se destaca por ser una metodología a través de la cual se efectúan pequeñas entregas al cliente de la funcionalidad del software final, haciendo partícipe a él en todo el proceso de desarrollo del producto para luego llegar a las pruebas de aceptación, dando su certificación [50].

Para el trabajo de titulación las fases de la metodología extrema se las trabajó de la siguiente manera:

✓ Planificación

En esta fase se realizó la recopilación de la información de los pacientes seleccionados para este estudio en el departamento de la Consulta Externa del Hospital con la entrevista y la revisión de trabajos similares se pudo hacer un análisis de la información recolectada y mediante ello complementar los requerimientos del software que se detallan en el Documento de Especificación de Requerimientos según el estándar IEEE 830. (Ver Anexo 4: Documento de especificación de Requerimientos).

✓ Diseño

En esta fase de la metodología se realizó el desarrollo de la arquitectura de la aplicación web (ver Anexo 6), en el cual se utilizó el modelo 4+1 propuesto por Kruchten donde en cada una de las vistas se propone un diagrama UML con los cuales la metodología de desarrollo se fortalece aún más.

✓ Desarrollo.

En esta fase se procedió a la codificación de la aplicación web, iniciando con la creación del proyecto Cardio en el framework django, posteriormente la

codificación de las vistas, finalmente con la programación de las métricas para la predicción del riesgo cardiovascular y el cálculo del perfil lipídico, todo esto de acuerdo con los requerimientos obtenidos en el Documento de Especificación de Requerimientos.

✓ **Pruebas.**

En esta fase de la metodología se desarrolló las pruebas de aceptación a los médicos especialistas de la Consulta Externa del Hospital Ceibos, los mismos que tuvieron una muy buena aceptación, de igual manera se realizaron las pruebas unitarias a las funcionalidades del sistema, estas pruebas se realizaron durante el transcurso del proyecto para así evitar futuros problemas similares. (ver Anexo 9)

5.4. Participantes.

EL presente trabajo de titulación fue desarrollado por Roger Alexander Torres Yaguana autor del mismo, conjuntamente con el director y tutor académico Ing. Francisco Javier Álvarez Pineda. De igual manera se tuvo la colaboración del Dr. Antonio Pozo García y el Dr. Lenin Torres Y. del departamento de Consulta Externa del Hospital General Norte de Guayaquil Ceibos.

6. Resultados

Para el desarrollo del trabajo de titulación se planteó definir las fases por cada objetivo planteado, siendo estas esenciales para la terminación favorable del presente trabajo. Así se enuncian a continuación los aspectos que fueron utilizados en el desarrollo del trabajo de titulación:

6.1. FASE UNO: Realizar una revisión bibliográfica de trabajos relacionados

6.1.1. Planificación de la literatura de Proyectos Relacionados

En la realización de esta fase, correspondiente a la revisión bibliográfica de trabajos relacionados con este proyecto, se adoptó el siguiente flujo de actividades para poder obtener la información relevante de la hipertensión, diabetes mellitus, su asociación con el colesterol y el riesgo cardiovascular, puntualizando las mejores técnicas de aprendizaje automático y sus resultados obtenidos; encontrando así la literatura más adecuada para la realización de este trabajo investigativo.

Justificación.- La grave situación de este tipo de enfermedades preexistentes como son la Diabetes Mellitus Tipo 2 y la Hipertensión Arterial Sistémica en nuestro medio hacen necesario que se utilicen herramientas tecnológicas que coadyuven a la parte médica a manejarlas mejor y con precisión para estas problemáticas médicas, ya que algunos casos los pacientes fallecieron porque nunca se llevó a cabo un control de su riesgo cardiovascular y sus niveles de colesterol, siendo este trabajo un aporte importante para el tratamiento y seguimiento de estas comorbilidades y con ello disminuir las potenciales complicaciones que podrían terminar en desenlaces irreversibles y/o fatales.

Recopilación. - La información se recogió de acuerdo a los criterios técnicos del especialista en Medicina Interna, siendo la diabetes, la hipertensión arterial, la obesidad, los mayores riesgos de eventos catastróficos en los pacientes seleccionados, obteniéndose en las bases de datos de las bibliotecas mundiales como son: PubMed, UpToDate, Science Direct, Scielo, Dialnet, Redalyc, Academic Search, IEEE.

Evaluación. - La relevancia de los artículos científicos seleccionados se basó en las cadenas de búsqueda y los criterios de calidad realizadas de las comorbilidades de las enfermedades cardiovasculares (Diabetes Mellitus e Hipertensión Arterial), por lo que los resultados obtenidos a través del aprendizaje automático fueron superiores a los métodos tradicionales de la medicina convencional.

Elaboración. - De acuerdo al problema de salud planteado en este Trabajo de Titulación; es decir, la correlación entre el perfil lipídico y el riesgo cardiovascular, los autores mostraron sus altos valores de precisión al momento de emplear el aprendizaje automático, aplicando la regresión lineal múltiple, para esta temática sanitaria, debido a la forma de la información de las variables utilizadas en dicho estudio y ser uno de los métodos más usados en problemáticas médicas. Y así esta tecnología se puede aplicar a otras enfermedades que aquejan a la sociedad actual para su respectivo control y prevención.

Para la selección de documentos de trabajos relacionados con este tipo de proyecto se escogieron las investigaciones relacionadas a la enfermedad cardiovascular (ECV), preferentemente a la Diabetes Mellitus tipo 2 e Hipertensión Arterial Sistémica, donde se buscó trabajos investigativos que utilicen las técnicas de machine learning de la inteligencia artificial para hacer predicción, información buscada en las bibliotecas de Google Scholar, Dialnet, Scielo, Redalyc, Academic Search, IEEE, PubMed, UpToDate; otorgando así la información necesaria para la consecución de este proyecto de investigación. (Ver Tabla 10.), capítulo 1 Sección dos.

Realizando un análisis de todos estos estudios las técnicas de aprendizaje automático, donde se encontró 2 técnicas de árboles de decisión, 3 de máquinas de soporte vectorial, 3 de redes bayesianas, 4 de regresión lineal y 2 de redes neuronales artificiales.

En base a las técnicas encontradas en el apartado anterior se ha optado por emplear la técnica de la regresión lineal múltiple por cuanto más investigaciones de problemas de salud relacionados se realizan con esta técnica que es la base a partir de la cual se construyeron los otros métodos de aprendizaje profundo, siendo así la herramienta más idónea para la consecución de este trabajo de titulación.

A continuación, se elaboró la siguiente tabla de selección de documentos de los trabajos con mejores resultados de rendimiento para el seguimiento de estas enfermedades, realizando un análisis de sus técnicas, las metodologías empleadas y los resultados obtenidos para la consecución de los objetivos planteados en dichos trabajos investigativos relacionados con el presente trabajo de titulación.

Tabla 9. Selección de Documentos

Código	Título	Términos	Año de publicación	Buscador
ES01	Análisis y diseño de un sistema para identificar signos de retinopatía hipertensiva a través de imágenes de retina, aplicando la tecnología de deep learning	Prototipo, inteligencia artificial, algoritmo	2018	Google académico
ES02	Modelo de minería de datos usando machine learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016	Datos, aprendizaje automático, machine learning	2016	Google académico
ES03	Aplicación de la Inteligencia Artificial en la Bioinformática, avances, definiciones y herramientas.	Bioinformática, predicción, técnicas.	2016	Google académico

ES04	Modelo tecnológico para el control y monitoreo del cuidado médico de enfermedades crónicas usando wearables de bajo costo	Salud, sistema, población.	2018	Academic Search
ES05	Construcción de un modelo de predicción para apoyo al diagnóstico de diabetes	BigML, aprendizaje supervisado, machine learning, sistema experto.	2018	Redalyc
ES06	Una guía conceptual para usar y entender Big Data en la investigación clínica	Análisis de datos, aprendizaje de máquina, inteligencia artificial, minería de datos.	2019	Google académico
ES07	Evaluación de triglicéridos y colesterol en personas con sobrepeso basada en la regresión lineal múltiple y el modelo de inteligencia artificial	Triglicéridos, colesterol, red neuronal, inteligencia artificial	2017	Google académico
ES08	Asociación de puntuación de patrón de perfusión mejorada con la gravedad de la diabetes tipo 2 usando la tubería de aprendizaje automático: estudio piloto	Diabetes mellitus, machine learning,	2019	Google académico

ES09	Antecedentes de adherencia a tratamiento en pacientes hipertensos de un centro de salud chileno	Hipertensión arterial, información sociodemográfica, regresión lineal múltiple	2017	Redalyc
ES10	Asociación de bifenilos policlorados en sangre y niveles de colesterol entre los Inuit Canadienses	Colesterol, triglicéridos en sangre PCB.	2018	Science Direct
ES11	La asociación prospectiva entre la inflamación y los síntomas depresivos en la diabetes tipo 2 estratificada por sexo	Diabetes, colesterol, modelo regresión lineal, síntomas.	2019	PubMed
ES12	Desarrollo de Biomarcadores basados en determinantes socioeconómicos y demográficos para el diagnóstico de Enfermedades Multifactoriales no transmisibles por medio de Redes Neuronales profundas	Diabetes mellitus, redes neuronales artificiales, diagnóstico asistido por computadora, caries.	2019	Google académico
ES13	Implementación de algoritmos de inteligencia artificial para la identificación de pacientes	Algoritmos de inteligencia artificial, diabetes,	2018	Google académico

	diabéticos utilizando los niveles de lípidos en sangre	colesterol, triglicéridos.		
ES14	Métodos de aprendizaje automático para la detección de anomalías	Outliers, detección de anomalías, aprendizaje automático.	2019	Google académico
ES15	Valoración de riesgo cardiovascular mediante modelos de clasificación	Clasificadores, clasificación estadística, árbol de regresión, regresión logística.	2019	Google académico
ES16	Sistema e-health para capturar, transmitir y almacenar el estado de la presión arterial y otros datos relevantes, con el fin de calcular la probabilidad de padecer hipertensión arterial.	E-health, inteligencia artificial, red de bayes, redes de datos, telecomunicaciones, sistema de captación, procesamiento de información, probabilidad.	2020	Google académico
ES17	Diseño de un prototipo web interactivo basado en técnicas de machine learning y razonamiento probabilístico para pronosticar mediante los	Diabetes, machine learning, razonamiento probabilístico, regresión logística	2019	Google académico

	factores de riesgo la enfermedad de diabetes.			
ES18	Modelo basado en aprendizaje de máquina supervisado, para el análisis de datos cardiovasculares y su aplicación en el diagnóstico y pronóstico médico mediante el uso de una página web dinámica.	Enfermedades cardiovasculares, diagnóstico, machine learning, naive bayes, redes bayesianas	2019	Google académico
ES19	Contribuciones de las técnicas machine learning a la cardiología. Predicción de reestenosis tras implante de stent coronario	Inteligencia artificial, machine learning, deep learning, cardiología	2019	Google académico
ES20	Extracción de información de Evoluciones Clínicas Digitales mediante técnicas de Machine Learning	Machine learning, extracción de información, factores de riesgo cardiovascular.	2019	Google académico
ES21	Desarrollo de una aplicación móvil para el monitoreo de pacientes con hipertensión arterial utilizando la tecnología ESP32	Aplicaciones móviles, módulo ESP32, hipertensión, telemedicina, monitoreo presión arterial	2020	Google académico

ES22	Minería de datos para la detección de factores de influencia en el test APGAR	Minería de datos, algoritmos de clasificación, test de Apgar	2019	Google académico
ES23	Diabetes Mellitus Tipo 2 y sus efectos sobre procesos de Cognición Social	Diabetes, cognición social, deterioro cognoscitivo	2019	Google académico
ES24	Estudio de la relación del esfuerzo de corte con la presión en aneurismas cerebrales y la predicción del riesgo de ruptura usando herramientas de inteligencia artificial basado en datos morfológicos, fluidodinámicos y estructurales	Inteligencia artificial, aneurismas, hipertensión	2019	Google académico
ES25	Desarrollo de una aplicación móvil (Appteroma) que permita estimar el riesgo de padecer Aterosclerosis utilizando algoritmos de Machine Learning	Aterosclerosis, aplicación móvil, machine learning.	2019	Google académico
ES26	Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí	Inteligencia artificial, machine learning, deep learning, cardiología	2019	Science Direct

ES27	Diseño y evaluación de un sistema de ayuda al diagnóstico para neuropatía diabética basado en la lectura de puntos de presión plantar y machine learning	Presión plantar, sistema de apoyo al diagnóstico, sensores, machine learning, support vector machine, neuropatía diabética	2019	Google académico
ES28	Modelos de aprendizaje profundo para la predicción de comportamiento sedentario futuro	Sedentarismo, enfermedades cardiovasculares, predicción, aprendizaje profundo	2019	Google académico
ES29	Análisis neutrosófico para el diagnóstico de la hipertensión arterial a partir de un sistema experto basados en casos	Sistema experto basado en casos, hipertensión arterial, técnicas estadísticas, neutrosófica, modelo de recomendación	2019	Google académico
ES30	Red neuronal para diagnóstico de hipertensión arterial	Clasificadores, perceptrón multicapa, árbol de decisión C4.5	2019	Google académico
ES31	Índice HOMA-IR como predictor de reducción de exceso de peso en pacientes con índice de masa corporal	Índice HOMA-IR, índice de masa corporal, diabetes mellitus tipo 2,	2020	Google académico

	IMC \geq 35 kg/m ² sometidos a gastrectomía vertical	hipertensión arterial, dislipidemia		
ES32	Métodos de aprendizaje supervisado para la predicción de diabetes: una revisión sistemática de literatura	Inteligencia artificial, aprendizaje supervisado, predicción de diabetes.	2019	Google académico
ES33	Determinación de cristatina c e identificación de la relación con la creatinina sérica en el diagnóstico de insuficiencia renal en pacientes con diabetes mellitus tipo 2	Cristatina C, creatinina, daño renal, filtración glomerular, diabetes mellitus	2017	Google académico
ES34	Una guía conceptual para usar y entender Big Data en la investigación clínica	Análisis estadístico de datos, aprendizaje de máquina, inteligencia artificial, minería de datos	2019	Science Direct
ES35	Un enfoque de aprendizaje automático para la estimación de diagramas de fase para mezclas de lípidos de tres componentes	Lípidos, colesterol, aprendizaje automático, redes neuronales, bosques aleatorios	2020	Science Direct

ES36	Clasificación de factores sociodemográficos, de comportamiento de salud, de prevención y ambientales en la predicción de la salud cardiovascular del vecindario: un enfoque de aprendizaje automático bayesiano	Comportamiento de salud, prevención, salud cardiovascular, barrio, aprendizaje automático	2020	Science Direct
ES37	Prevención del riesgo cardiovascular/accidente cerebrovascular: un nuevo marco de aprendizaje automático que integra fenotipos basados en imágenes de ultrasonido carotídeo y sus armónicos con factores de riesgo convencionales	Aterosclerosis, factores de riesgo convencionales, covariables, carótida, ultrasonido, fenotipos basados en imágenes	2020	Science Direct
ES38	Un modelo predictivo preciso y dinámico para un sistema inteligente M-Health que utiliza aprendizaje automático	Aprendizaje automático, predictivo, modelos, clasificación, SVM, árbol de decisión, precisión	2020	Science Direct
ES39	Como el aprendizaje automático transformará la biomedicina	Aprendizaje automático,	2020	Science Direct

		biomedicina, diagnóstico		
ES40	Imagen confocal asistida por aprendizaje automático de los sitios intracelulares de formación y almacenamiento de triglicéridos y ésteres de colesterol	Lípidos, imágenes metabólicas, aprendizaje automático, segmentación, medicina personalizada	2020	Science Direct
ES41	Una revisión de las técnicas de aprendizaje automático en fotopletismografía para la medición no invasiva sin manguito de la presión arterial	Fotopletismografía, presión arterial, medición, aprendizaje automático	2020	Science Direct
ES42	Investigación sobre el efecto del número de Womersley, las funciones de ECG y PPG para la estimación de la presión arterial sin manguito mediante el aprendizaje	Presión arterial continua sin manguito, bosque aleatorio, parámetro de Womersley, funciones de ECG, algoritmo genético	2020	Science Direct
ES43	Un nuevo método de estimación de la presión arterial basado en la clasificación de formas de ondas oscilométricas	Forma de onda oscilométrica, clasificación, estimación de la presión arterial,	2020	Science Direct

	mediante métodos de aprendizaje automático	dispositivo médico		
ES44	Aprendizaje automático en el modelado de flujos cardiovasculares: predicción de la presión arterial a partir de datos de resonancia magnética de flujo 4D no invasivos utilizando redes neuronales informadas por la física	Redes neuronales profundas, modelado de flujo sanguíneo, propagación de ondas de pulso, modelado basado en datos, diagnóstico no invasivo	2020	Science Direct
ES45	Aprendizaje de propiedades físicas en escenas visuales complejas: una máquina inteligente para percibir la dinámica del flujo sanguíneo a partir de imágenes de angiografía por TC estática	Aprendizaje automático, RNN estructurado en árbol, reserva de flujo fraccional, angiografía por TC	2020	Science Direct
ES46	Enfoque novedoso que combina aprendizaje máquina y modelado de dinámica de sistemas para la predicción del riesgo de eventos cardiovasculares secundarios y el efecto de las intervenciones	Riesgo cardiovascular, aprendizaje automático, procedimiento	2020	Science Direct
ES47	Requisitos propuestos para la evaluación del	Inteligencia artificial, lista de	2020	Science Direct

	aprendizaje automático relacionado con imágenes cardiovasculares (PRIME): una lista de verificación: revisada por el Consejo de innovación en el cuidado de la salud del American College of Cardiology	verificación aprendizaje automático, salud digital		
ES48	Integración de la puntuación STOP-BANG y los datos clínicos para predecir eventos cardiovasculares después de un infarto: un estudio de aprendizaje automático	Infarto agudo de miocardio, eventos cardiovasculares, aprendizaje automático, puntaje STOP-BANG	2020	Science Direct
ES49	Uso del aprendizaje automático para predecir eventos cardiovasculares de un año en pacientes con miocardiopatía dilatada grave	Miocardiopatía dilatada grave, valor pronóstico, aprendizaje automático	2019	Science Direct
ES50	Inteligencia artificial y aprendizaje automático en la atención de la salud cardiovascular	Inteligencia artificial, aprendizaje automático, salud cardiovascular	2020	Science Direct

ES51	Detección de la gravedad de los trastornos cardiovasculares mediante el uso de características anatómicas del miocardio basado en un enfoque de máquina de aprendizaje extremo optimizado	Trastorno cardiovascular, imágenes de resonancia magnética, miocardio, máquina de aprendizaje extrema, optimización de mariposa	2020	Science Direct
ES52	Un enfoque basado en el aprendizaje automático para predecir el brote de enfermedades cardiovasculares en pacientes en diálisis	Aprendizaje automático, resultados cardiovasculares, ESRD, pronóstico	2019	Science Direct
ES53	Técnicas de aprendizaje automático para predecir los resultados cardiovasculares en el Hospital en pacientes ancianos que presentan infarto de miocardio agudo	Aprendizaje automático, predicción, infarto de miocardio	2020	Science Direct
ES54	Poner en marcha el aprendizaje automático: aplicaciones en imágenes cardiovasculares	Imágenes cardíacas, aprendizaje automático, predicciones	2020	Science Direct
ES55	Usos y oportunidades del aprendizaje automático en la	Aprendizaje automático, redes	2020	Science Direct

	investigación de la hipertensión	neuronales profundas, hipertensión		
ES56	Predicción dinámica de la mortalidad mediante técnicas de aprendizaje automático para casos cardiovasculares agudos	Aprendizaje automático, síndrome coronario agudo, cardiología.	2018	Science Direct
ES57	Aprovechamiento de las técnicas de aprendizaje automático para pronosticar el pronóstico del paciente después de una intervención coronaria percutánea	Aprendizaje automático, intervención coronaria percutánea, rehospitalización	2019	Science Direct
ES58	Aprovechar los algoritmos de aprendizaje automático interpretables para predecir los resultados posoperatorios de los pacientes en dispositivos móviles	Aprendizaje automático, predicción, dispositivos móviles	2020	Science Direct
ES59	El aprendizaje automático mejora la predicción del riesgo de mortalidad después de la cirugía cardíaca: revisión sistemática y metaanálisis	Modelo de riesgo, mortalidad, aprendizaje automático, regresión logística, metaanálisis	2020	Science Direct

ES60	Análisis de redes y aprendizaje automático para el modelado predictivo del riesgo de enfermedad cardiovascular en pacientes con diabetes tipo 2	Datos administrativos, comorbilidad, enfermedad cardiovascular, diabetes tipo 2, aprendizaje automático	2020	Science Direct
ES61	Aprendizaje automático para predecir el riesgo a largo plazo de infarto de miocardio y muerte cardíaca en función del riesgo clínico, calcio coronario y tejido adiposo epicárdico: un estudio prospectivo	Puntuación de calcio coronario, tejido adiposo epicárdico, aprendizaje automático, infarto de miocardio, muerte cardíaca	2019	PubMed
ES62	El aprendizaje automático revela que los esfingolípidos séricos son biomarcadores independientes del colesterol de la enfermedad de las arterias coronarias	Enfermedad cardiovascular, lipoproteínas, metabolismo, Biología vascular	2020	PubMed
ES63	La lipidómica de espectrometría de masas no dirigida identifica la correlación entre las esfingomielinas séricas y el colesterol plasmático	Colesterol HDL, lipidómica, lipoproteínas, esfingolípidos	2019	PubMed

ES64	Valor predictivo de la proteómica dirigida para la morfología de la placa coronaria en pacientes con sospecha de enfermedad arterial coronaria	Biomarcadores, arteriopatía coronaria, angiografía por tomografía computarizada coronaria, proteómica, evaluación de riesgos.	2019	PubMed
ES65	El lipidoma cardíaco en modelos de enfermedad cardiovascular	Enfermedad cardiovascular, insuficiencia cardíaca, lípidos, espectrometría de masas; infarto de miocardio, obesidad.	2020	PubMed
ES66	Predecir las complicaciones de la diabetes mellitus mediante algoritmos avanzados de aprendizaje automático	Modelos RNN, aprendizaje profundo, diabetes mellitus, aprendizaje automático.	2020	PubMed
ES67	Predecir el riesgo a 10 años de complicaciones en los órganos terminales de la diabetes tipo 2 con y sin cirugía metabólica: un	Predicción, diabetes tipo 2, aprendizaje automático	2020	PubMed

	enfoque de aprendizaje automático			
ES68	Explorando factores de riesgo de insuficiencia cardíaca a través del aprendizaje automático	Infarto de miocardio, aprendizaje automático	2019	Google académico
ES69	Los niveles séricos de ácido úrico y colesterol unido a lipoproteínas de baja densidad son predictores independientes de enfermedad arterial coronaria en pacientes indios asiáticos con diabetes mellitus tipo 2	Arteriopatía coronaria, enfermedad cardiovascular, diabetes mellitus tipo 2.	2016	PubMed
ES70	Partículas similares a remanentes y enfermedad de las arterias coronarias en la hipercolesterolemia familiar.	Hipercolesterolemia familiar, receptor de LDL, triglicéridos, regresión logística.	2018	PubMed

6.1.2. Análisis bibliográfico.

En el presente apartado se va a describir el análisis de un estudio relacionado con el presente trabajo de titulación. Y los restantes estudios están descritos en el Anexo 2.

- ✓ ES60: En este estudio: “Análisis de redes y aprendizaje automático para el modelado predictivo del riesgo de enfermedad cardiovascular en pacientes con diabetes tipo 2”, se propuso un modelo de predicción de riesgos de un conjunto de datos administrativos de los fondos sanitarios privados en Australia utilizando aprendizaje automático ML para evaluar el riesgo de enfermedades cardiovasculares ECV en

pacientes con diabetes tipo 2, utilizando dos cohortes: uno de pacientes con diabetes tipo 2 y enfermedad cardiovascular y otro con solo pacientes con diabetes tipo 2. Se generaron dos redes de enfermedades de referencia a partir de las dos cohortes, luego se generó una red de enfermedades final a partir de dos redes de enfermedades de referencia mediante normalización. Se extrajo algunas características como la prevalencia de comorbilidades, patrones de transición y pertenencia a agrupaciones. Estos factores de riesgo se usaron para desarrollar seis modelos de predicción de ML. La precisión de los clasificadores varió del 79% al 88% mostrando el poder de predicción de riesgos basados en la red y el ML de los datos administrativos. Este modelo se debería utilizar para la gestión de pacientes con alto riesgo de desarrollar enfermedades crónicas [51].

6.1.3. Reporte de la Revisión Bibliográfica

Luego del análisis de los artículos científicos relacionados con el presente proyecto de titulación se llegó a determinar que dentro de las diferentes técnicas de aprendizaje automático de la Inteligencia Artificial utilizadas para los estudios referentes a este tema específico de las enfermedades cardiovasculares; es decir, el riesgo cardiovascular y perfil lipídico, la regresión lineal múltiple es la técnica más idónea para obtener mejores resultados en función de que se está trabajando con información numérica y es el método más utilizado para temáticas médicas, donde el conjunto de variables usadas permitirá la inclusión de un mayor número de las mismas, hablamos de presión arterial diastólica, colesterol, edad, índice de masa corporal, entre otras. Así, se llegó a escoger conjuntamente la técnica del aprendizaje automático que es la regresión lineal múltiple para el seguimiento del perfil lipídico y su relación con el riesgo cardiovascular, por las características y los valores numéricos de la información de los pacientes con diabetes mellitus tipo 2 e hipertensión arterial sistémica y relacionados, registrados en la casa de salud seleccionada, concretamente el Hospital General del Norte de Guayaquil “Los Ceibos”.

Conjuntamente con el médico especialista en Medicina Interna, soporte de la parte médica para esta investigación, se corroboraron los 14 artículos científicos que guardan relación directa con el tema de colesterol-diabetes e hipertensión, puntualizando que no se va a diagnosticar la enfermedad de diabetes ni la hipertensión, sino se va a

correlacionar a pacientes que ya tienen la enfermedad con el colesterol y su riesgo de complicaciones a mediano y largo plazo.

6.2. FASE DOS: Recopilación y limpieza de datos para el perfil lipídico y el riesgo cardiovascular

Para la realización de esta fase, referente a la recopilación y limpieza de la información requerida de los pacientes seleccionados para esta temática médica de investigación, se la desarrolló de acuerdo a las siguientes actividades, que a continuación se muestran.

6.2.1. Análisis de herramientas informáticas para el proyecto.

Para la realización de este apartado, se seleccionó las siguientes herramientas informáticas, las cuales sirvieron para la construcción de la aplicación web con aprendizaje automático supervisado.

Django

Django, promovió el desarrollo diligente del trabajo de titulación con el ahorro significativo de tiempo para su ejecución, previendo de hacer el trabajo eficiente con el conocimiento suficiente. Este framework de alto nivel escrito en Python permitió desarrollar múltiples funcionales de la aplicación web, como los modelos, vistas, formatos [52].

Dentro de las ventajas y desventajas en el aplicativo, se obtuvieron las siguientes:

- Ventajas:
 - ✓ Ofreció mayor seguridad a la aplicación y facilitó su mantenimiento, haciendo uso del Modelo Vista Plantilla (MVT), de manera eficiente.
 - ✓ El empleo del objeto mapa relacional (ORM), propio de Django, facilitó el trabajo de la programación cuando se elaboró la capa del modelo de la aplicación.
 - ✓ Fue sencilla la inclusión en la aplicación, para la personalización y también para el caso de la herencia.
- Desventajas:

- ✓ Algunas veces se requirió incrustar código SQL puro para realizar ciertas consultas complejas.
- ✓ Para la manipulación de configuraciones más específicas, resultó complejo en la codificación ya que se tenía una categoría de programador junior.

Estas fueron algunas de las ventajas y desventajas para mencionar del framework de desarrollo Django.

Python

El lenguaje Python fue muy versátil y a la vez conveniente para la automatización de los procesos realizados de acuerdo a la metodología de programación ágil XP, en el sistema operativo empleado [53].

- Ventajas:
 - ✓ Por manejar una estructura muy simplificada en su desarrollo fue más rápido su uso en el aplicativo.
 - ✓ Siendo un lenguaje multiplataforma, se pudo integrar con otros dispositivos.
 - ✓ Sus herramientas de seguridad otorgaron las respectivas mejoras para la aplicación web.
 - ✓ La gran comunidad de programadores permitió dar solución a la mayoría de los problemas presentados en el transcurso del proyecto.
- Desventajas:
 - ✓ Los programas interpretados son más lentos que los compilados.
 - ✓ No se programó en un bajo nivel, ya que sus instrucciones ejercen control directo sobre el hardware.
 - ✓ Al ser un lenguaje interpretado, generó algunos errores de escritura.

Estas fueron algunas de las ventajas y desventajas del lenguaje de programación Python por mencionar en este trabajo de titulación.

Librerías de Python:

- Scikit-learn

Esta librería de Machine Learning gratuita para Python, fue indispensable para realizar la predicción, ya que contaba con amplias bibliotecas numéricas y científicas de Python

y su documentación ayudó muchísimo en el trabajo. Se pudo implementar modelos de aprendizaje automático de manera sencilla [54], [55].

- **Pandas**

Esta librería de análisis de datos con Python, incluyó herramientas para lectura y escritura de datos. Sirvió para el manejo integrado de datos faltantes. Además, contribuyó de manera prominente en el ámbito de la salud humana, analizado en este trabajo de titulación desarrollado [55].

- **NumPy.**

Numerical Python, fue el paquete principal para hacer computación científica en Python que proporcionó grandes estructuras de datos, con un poderoso objeto de matriz N-dimensional. Se observó que Numpy tenía funcionalidades muy parecidas a otros lenguajes de programación como Matlab [55], [56].

Aprendizaje Automático

El aprendizaje representó la razón de ser de este trabajo de titulación, haciendo uso de recursos básicos para el conocimiento necesario para el mismo, ya que es una parte primordial de la Inteligencia Artificial que diseña algoritmos de computadora para aprender a través de la experimentación y así hacerlo por sí mismo. El aprendizaje automático permitió usar las ventajas que posee en esta área como es la medicina humana y en este tema en particular del perfil lipídico y el riesgo cardiovascular de los pacientes diabéticos tipo 2 e hipertensos y aprovechar sus funcionalidades. [57].

Además para el aprendizaje automático, se realizó el tratamiento minucioso de los datos recopilados de los pacientes para este proyecto investigativo desde la institución hospitalaria “Ceibos”, como parte esencial de la computación aplicada [58] [59].

Aprendizaje automático supervisado

Este tipo de aprendizaje específico dedujo una función principal a partir de un conjunto de datos de entrenamiento, estos estaban ordenados por pares de objetos y la salida en una función que etiquetaba objetos relacionados con las entradas, donde la computadora aprendía de una multitud de datos ingresados y a partir de ahí poder hacer el resto de

cálculos necesarios para que no haya que volver a ingresar más información. Este tipo de aprendizaje supervisado en el proyecto se empleó para pronosticar el riesgo cardiovascular a través de las variables predictoras seleccionadas bajo el acompañamiento de la parte médica. Se debe mencionar que los tipos de algoritmos de aprendizaje supervisado más utilizados son: k-vecinos más cercanos, la regresión lineal, la regresión logística, los vectores de máquina de soporte, árboles de decisión, bosques aleatorios y las redes neuronales, para trabajos afines a esta investigación realizada [60].

Regresión Lineal

Esta técnica estadística sirvió para establecer la relación que existía entre dos o más variables en la información recopilada de los pacientes seleccionados en el presente trabajo. Primero empleando solamente dos variables, una dependiente y otra independiente para obtener una regresión lineal simple y posterior a ello emplear más de dos variables, una predicha y varias predictoras donde se obtuvo una regresión lineal múltiple, con los registros de pacientes con las enfermedades preestablecidas de la diabetes mellitus tipo 2 e hipertensión arterial sistémica [61].

Se debe mencionar que existe la posibilidad de poder considerar dos valores como variables de entradas o predictoras en la regresión lineal múltiple, como se visualiza en la Figura 7, a continuación, en un modelo de regresión con retrodispersión, a manera de ejemplo.

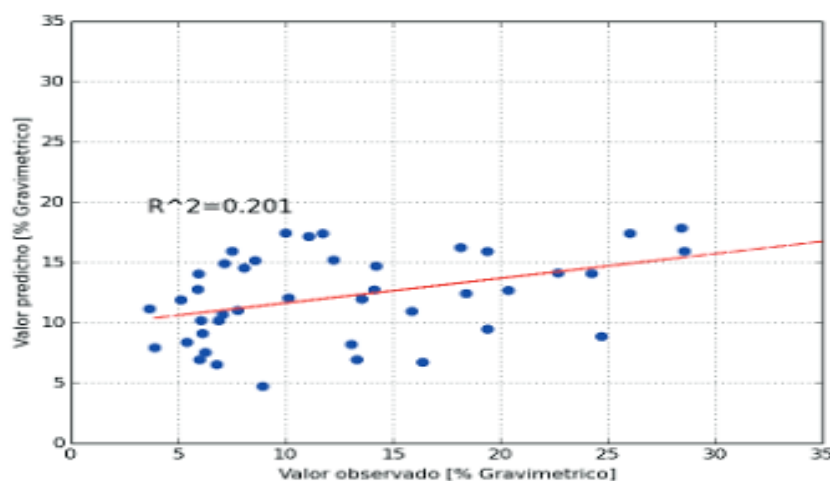


Figura 7. Modelo de regresión lineal múltiple con retrodispersión

En este modelo de regresión lineal múltiple con retrodispersión desde la parte matemática, la regresión lineal se expresaría de la siguiente forma:

$$Y = f(x_1, x_2, \dots) + \varepsilon$$

Dentro de los modelos de regresión que resaltaron para este modelo fueron: el modelo de regresión lineal y el modelo de regresión logística. Existen otros modelos de regresión utilizados como la regresión no lineal o la regresión ordinal que estiman una serie de modelos matemáticos que pueden ajustarse mejor que un modelo lineal [62].

Importancia de la regresión lineal en el proyecto de titulación

Esta técnica fue muy útil para los procesos de inferencias de algún factor que se desconocía su resultado, en la cual se tenía información de una variable independiente y el comportamiento de otra variable que dependía de la primera. En este proyecto investigativo la importancia radicó en la utilización de la regresión lineal para esta temática médica, en donde el especialista en medicina interna a través de este aplicativo informático, a través de la predicción del riesgo cardiovascular pueda tener los valores necesarios para su respectiva toma de decisiones en cuanto al control, tratamiento y seguimiento del paciente analizado y mejorar su calidad de vida [63].

Regresión lineal múltiple

A través de esta técnica se examinó las relaciones lineales entre una respuesta continua y dos o más predictores. Para el presente trabajo de titulación, referente al modelo predictivo del riesgo cardiovascular las variables independientes empleadas fueron: el sexo, la edad, diabetes, el colesterol total, la lipoproteína de alta densidad, la presión arterial sistólica, fumador y el tratamiento antihipertensivo y la variable dependiente fue: riesgo. Respecto al resultado del modelo se lo obtuvo de acuerdo a un porcentaje del 0 al 100% con su respectivo rango de etiquetado en bajo, moderado, alto y muy alto riesgo. (Ver sección 4.2, subsección 4.2.2)

La ecuación es de la forma:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_l x_l + \varepsilon$$

Donde la ecuación de regresión lineal múltiple comienza estimando los valores de los coeficientes β_k del modelo de regresión, minimizando las diferencias entre los valores observados y los pronosticados [59].

Supuestos del modelo de regresión lineal

Se tomará en cuenta que para el modelo de regresión lineal empleado se debería cumplir con los siguientes supuestos:

- ✓ Linealidad: la relación de las variables será lineal.
- ✓ Independencia: los errores en la medición de variables explicativas son independientes.
- ✓ Homocedasticidad: los errores tendrán varianza constante.
- ✓ Normalidad: las variables seguirán una ley normal.
- ✓ No colinealidad: las variables independientes no estarán correlacionadas.

En este trabajo, la selección de la variable dependiente, las variables explicativas, la verificación de los supuestos del modelo, la interpretación del modelo y la bondad del modelo fueron los pasos necesarios para la aplicación de la técnica de la regresión en el modelo predictivo del riesgo cardiovascular del paciente [64].

6.2.1.1. Elaboración del documento de los Requerimientos Funcionales y no Funcionales de Software

Para este apartado se tuvo como finalidad definir las especificaciones funcionales y no funcionales para el desarrollo de este sitio web que a la vez permitió dar un seguimiento del perfil lipídico y el riesgo cardiovascular de los pacientes diabéticos e hipertensos; siguiendo los cánones de la Metodología XP, se detalló los requerimientos para el aplicativo, así:

Requerimientos Funcionales

A continuación, a través del análisis se llegó a obtener los siguientes requerimientos funcionales y no funcionales del proyecto. Se debe resaltar que algunas funcionalidades

no se pudieron concretar por cuanto se tuvo ciertas restricciones por parte de las normativas de la entidad de salud (Ver Anexo 4)

Tabla 10. Requerimientos Funcionales

Código	Descripción	Categoría
RF01	Autenticar de usuario	Evidente
RF02	Cargar los datos cuantitativos de laboratorio del paciente	Evidente
RF03	Limpiar los datos cuantitativos del conjunto de pacientes	Evidente
RF04	Mostrar los datos normalizados en base a métricas requeridas del sistema	Evidente
RF05	Ingresar datos para predecir el riesgo cardiovascular	Evidente
RF06	Mostrar el porcentaje del riesgo cardiovascular del paciente	Evidente
RF07	Mostrar gráficos estadísticos de los datos cargados en el sistema	Evidente
RF08	Ingresar datos para calcular el perfil lipídico	Evidente
RF09	Mostrar el nivel del perfil lipídico del paciente	Evidente
RF10	Enviar mediante correo los resultados obtenidos del perfil lipídico	Evidente

Requerimientos No Funcionales

Se emplearon los siguientes requerimientos no funcionales porque estaban acordes con la consecución del sitio web desarrollado para el seguimiento de este problema de salud a través de la aplicación informática.

Tabla 11. Requerimientos No Funcionales

Requisito	Descripción
Look and Feel	El sistema tendrá una interfaz amigable con el usuario.
Seguridad	El sistema mantendrá la información de forma segura
Disponibilidad	El sistema deberá dar confiabilidad continua durante las horas laborables.
Robustez	El sistema deberá ser capaz de manejar grandes volúmenes de datos de modo seguro y rápido
Confidencialidad	La información de los pacientes será manipulada de forma segura con fines netamente académicos
Desarrollado con una metodología	El sistema será desarrollado con la metodología de desarrollo ágil Programación Extrema

6.2.1.2. Diseño de la Arquitectura de software y requerimientos del sistema

Para el diseño de la arquitectura se utiliza el modelo arquitectónico 4+1 de Kruchten, porque utiliza cinco vistas simultaneas, cada una de las ellas aborda un conjunto específico de preocupaciones mediante el uso del Lenguaje de Modelamiento Unificado (UML), [25].

Vista general de la aplicación web

La vista general de una aplicación web es muy importante ya que nos permite observar cada uno de los componentes de la aplicación y cómo interactúan entre sí para el correcto funcionamiento del sistema. Como muestra la figura 8, el usuario hace la petición al navegador y este lo envía al servidor, el cual carga el modelo de inteligencia artificial requerido desde su motor de base de datos Sqlite si el modelo ya está creado, sino lo crea a partir de los archivos estáticos csv y finalmente devuelve la respuesta al usuario final.

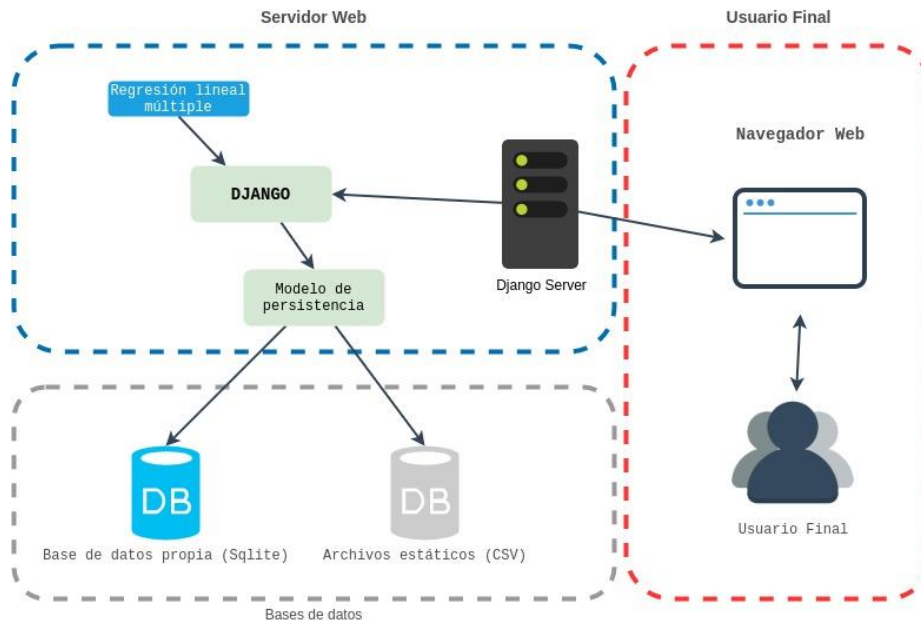


Figura 8. Modelo de Arquitectura MVT (Fuente Propia)

Arquitectura de Software

En esta sección se muestra los diagramas UML que componen la arquitectura de la aplicación web según el modelo 4+1 de Kruchten, en este modelo se proponen cuatro vistas (lógica, desarrollo, procesos y física) y una vista adicional (escenario) que es utilizada para vincular a las demás.

En la tabla VIII se muestra las vistas que se tomaron en cuenta para el desarrollo de la aplicación web.

Tabla 12: Representación de las Vistas del Modelo 4+1

Vistas	UML
Vista de Escenarios	Casos de Uso
Vista Lógica	Diagrama de Clases
Vista Física	Diagrama de Despliegue
Vista de Desarrollo	Diagrama de Componentes
Vista de Procesos	Diagrama de Actividad

Vista de Escenarios

Esta vista hace referencia al diagrama de casos de uso, donde se puede identificar los roles de usuarios y las actividades que pueden realizar cada uno de ellos, este diagrama sirve para especificar la comunicación y el comportamiento de la aplicación web mediante su interacción con los usuarios.

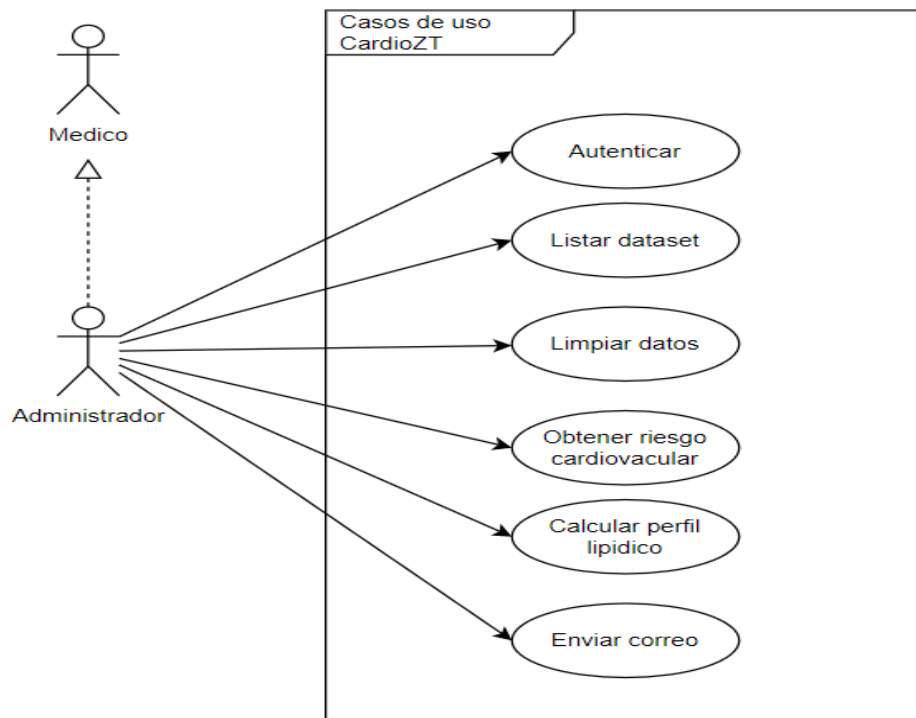


Figura 9: Diagrama de caso de uso (Fuente Propia)

Vista Lógica

En esta vista se muestra el diagrama de clases, que nos sirve para representar gráficamente y de manera estática la estructura general de la aplicación, en el que se muestra cada una de las clases existentes y sus interacciones.

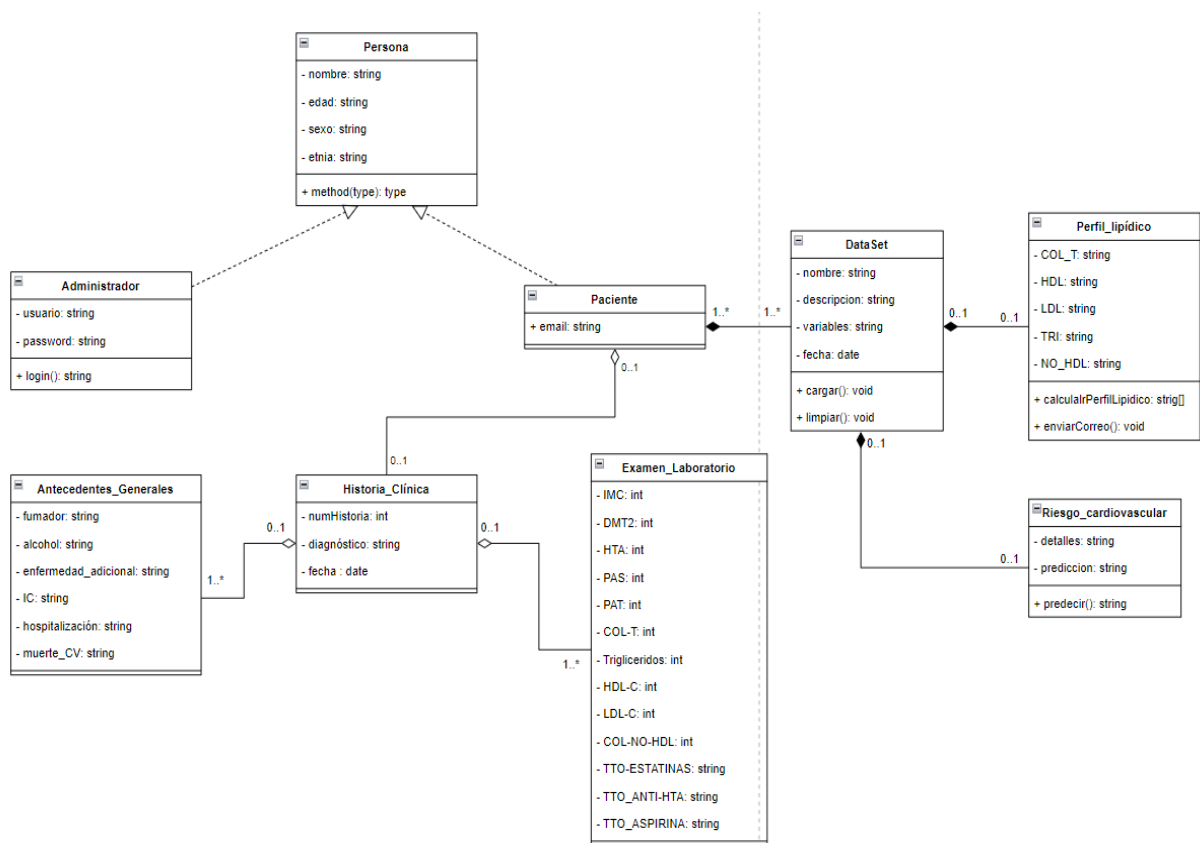


Figura 10. Diagrama de Clase

Vista Física

En esta sección se presenta la arquitectura física de la aplicación web mediante el diagrama de despliegue tal como se muestra en la figura 11, ya que nos permite representar la forma en que la aplicación se ejecuta sobre cada uno de los nodos (hardware), en este caso el usuario final, el computador, el servidor web y la base de datos, de igual forma nos muestra como los elementos de software se trazan en estos nodos y sus conexiones físicas entre estos elementos. (Ver Anexo 6).

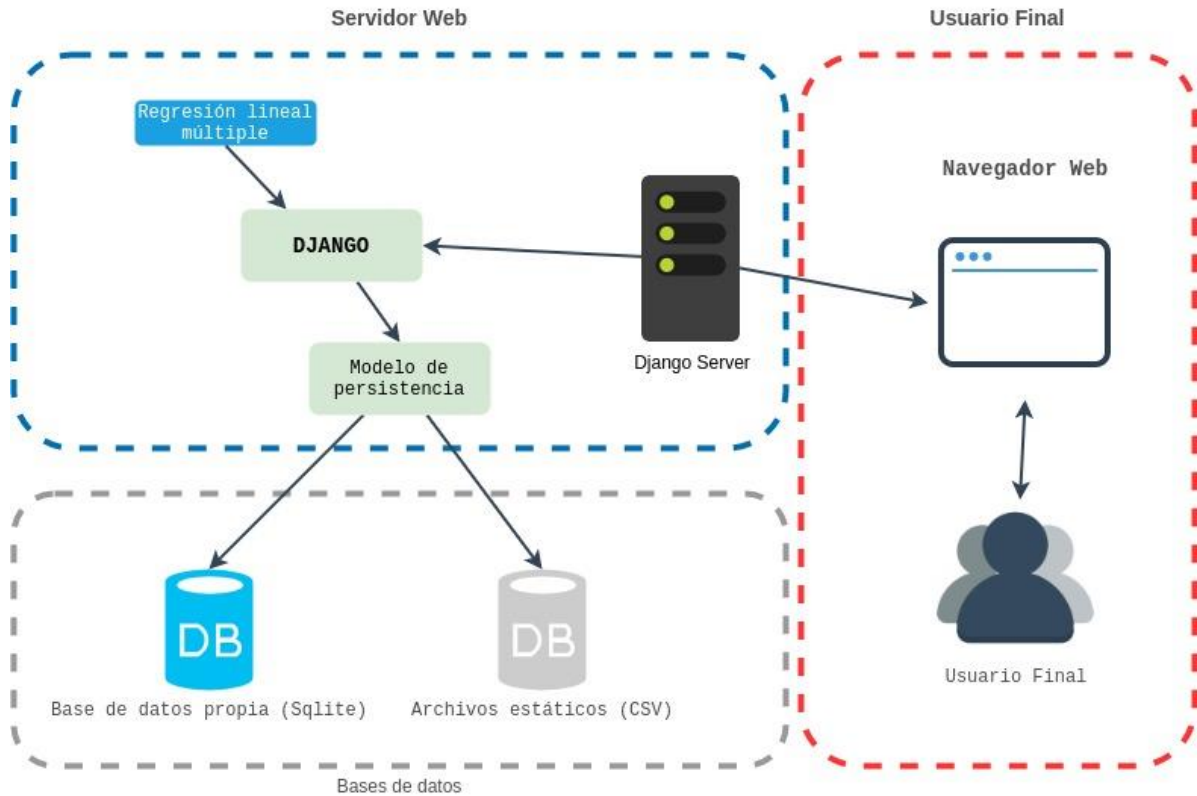


Figura 11: Diagrama de despliegue (Fuente Propia)

Vista de Desarrollo

En esta vista se muestra la forma de cómo están organizados los distintos componentes de la aplicación web en un ambiente de desarrollo de software, ya que cada uno de los módulos se empaqueta en partes pequeñas para facilitar su desarrollo, que a su vez apoya la evaluación de costos, la planificación, el monitoreo de progreso del proyecto y también como base para analizar el reúso, portabilidad y seguridad. Para la representación de esta vista se diseñó el siguiente diagrama de componentes.

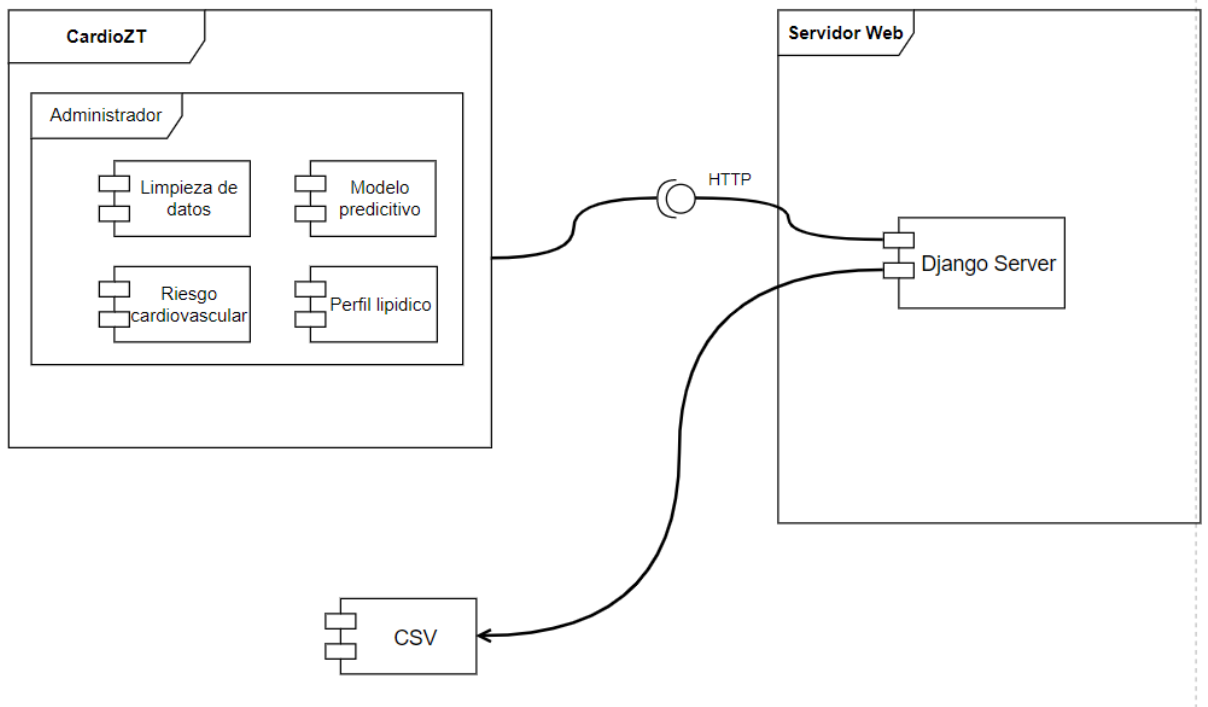


Figura 12: Diagrama de componentes (Fuente Propia)

Vista de Procesos

En esta vista se desarrollan los diagramas de actividad donde se puede observar el flujo de un proceso que es realizado mediante la interacción del usuario y la aplicación web desde el inicio hasta su finalización. A continuación, se muestran los principales diagramas de actividad de la aplicación web. (Ver Anexo 6).

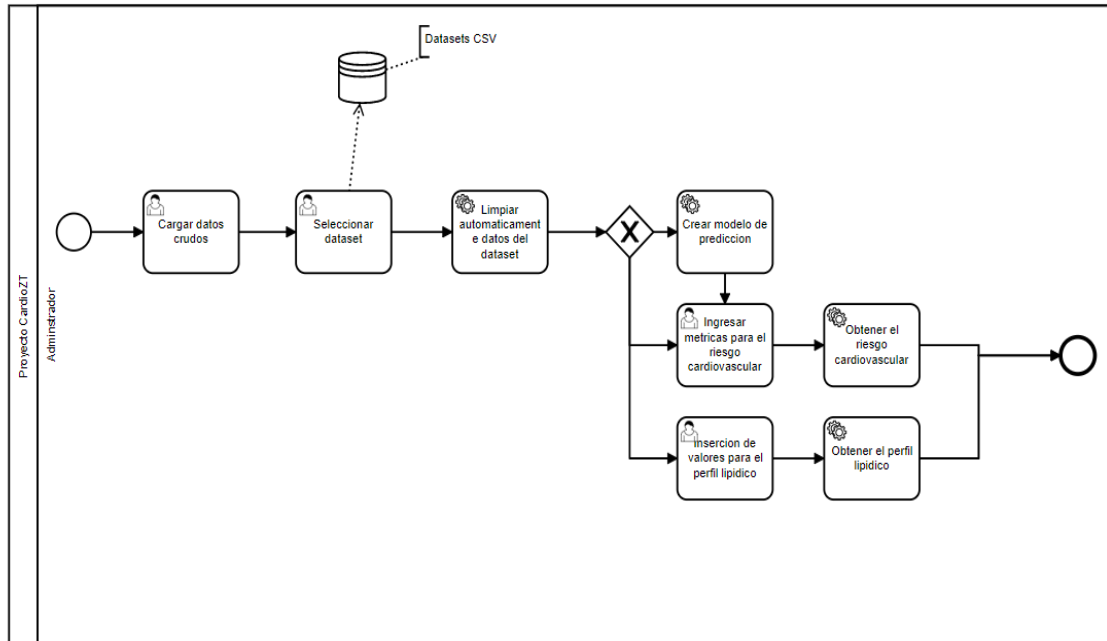


Figura 13: Diagrama de Actividad: Predicción del riesgo cardiovascular y el cálculo del perfil lipídico (Fuente Propia)

Resumen de Historias de Usuario.

Este es el resumen de las historias de usuario del sitio web, puntualizando que no es un sistema de gestión hospitalario. (Ver Anexo 7)

Tabla 13: Tabla de Historias De Usuario

N°	Nombre	Prioridad	Riesgo	Iteración
1	Autenticar usuario.	Alta	Medio	1
2	Cargar los datos cuantitativos de laboratorio del paciente.	Alta	Bajo	1
3	Limpiar los datos cuantitativos del conjunto total de pacientes	Media	Medio	1
4	Mostrar los datos normalizados en base a métricas requeridas del sistema	Media	Bajo	1
5	Ingresar datos para predecir el riesgo cardiovascular	Alta	Medio	1
6	Mostrar el porcentaje del riesgo cardiovascular del paciente	Alta	Medio	2
7	Mostrar gráficos estadísticos de los datos cargados en el sistema	Baja	Bajo	2
8	Ingresar datos para calcular el perfil lipídico	Media	Medio	2
9	Mostrar el nivel del perfil lipídico del paciente	Alta	Alto	2
10	Enviar por correo los resultados del perfil lipídico.	Media	Medio	3

6.2.2. Recopilación de la información de los pacientes

Para la recopilación de la información sanitaria, de acuerdo a esta problemática de salud se realizó el siguiente proceso, que se detalla a continuación:

Primeramente, se presentó una solicitud del Trabajo de Titulación al departamento de Docencia del Hospital General del Norte de Guayaquil “Los Ceibos”, el cual fue aprobado en un lapso de tiempo determinado para su posterior desarrollo. Con ello se tuvo un acceso manual al Sistema AS400 del IESS, donde se recopiló cada uno de los registros de la información de los pacientes seleccionados, a partir del año 2017, llegando a obtener un total de 5623 pacientes, con antecedentes de diabetes e hipertensión arterial y similares, donde se reclasificó en varios datasets (excels) aplicando criterios de inclusión: pacientes con diabetes mellitus, con hipertensión arterial, con obesidad, con insuficiencia cardíaca, con hospitalización y muerte cardiovascular. Los criterios de exclusión: pacientes menores de 19 años, enfermedad inflamatoria muscular y hepatopatía crónica CHILD-PUGH SCORE C. Luego se llegó a seleccionar conjuntamente con la parte médica y de acuerdo a la naturaleza de los datos de los pacientes relevantes, donde el algoritmo de aprendizaje automático supervisado: regresión lineal múltiple fue el más adecuado para trabajar, basándose en los criterios de los trabajos relacionados a este proyecto de titulación y la dirección del tutor académico, con respecto a las patologías crónicas de las Enfermedades Cardiovasculares en fuentes bibliográficas como Dialnet, PubMed, Google Scholar, Scopus, IEEE, ResearchGate para la elaboración del marco teórico. Los resultados fueron revalorados con el equipo de trabajo.

6.2.2.1. Estructura del Dataset

Acorde al previo análisis de los datos recogidos sobre los proyectos de predicción y aplicando los respectivos criterios médicos del especialista en Medicina Interna y que se encontraron con la fecha a partir del año 2017, se pudieron visualizar los datos. La información se la consolidó con las variables definidas a continuación:

- ✓ Período
- ✓ Dependencia
- ✓ Nombre del médico
- ✓ Historia clínica

- ✓ Identificación
- ✓ Nombre paciente
- ✓ Sexo
- ✓ Fecha de nacimiento
- ✓ Edad
- ✓ Fecha de atención
- ✓ Código de diagnóstico presuntivo
- ✓ Código de diagnóstico definitivo
- ✓ Colesterol total
- ✓ Triglicéridos
- ✓ Lipoproteína de alta densidad
- ✓ Lipoproteína de baja densidad
- ✓ Colesterol de No Lipoproteína de alta densidad
- ✓ Presión arterial sistólica
- ✓ Presión arterial diastólica
- ✓ Índice de masa corporal
- ✓ Fumador
- ✓ Alcohol
- ✓ Tratamiento con estatinas
- ✓ Tratamiento antihipertensivo
- ✓ Tratamiento con aspirina
- ✓ Insuficiencia Cardíaca
- ✓ Hospitalización
- ✓ Muerte de causa cardiovascular

De las variables de la información recopilada con las que se decidió emplear las siguientes, entendiendo su significado:

Acorde al análisis realizado con este trabajo de investigación para el riesgo cardiovascular y perfil lipídico se definió las variables adicionales que complementaron a los datos para poder tener un modelo de predicción de éxito fiable. Las variables que fueron seleccionadas finalmente, son:

- ✓ Colesterol total: cantidad total del colesterol en la sangre.
- ✓ Triglicéridos: tipo de grasa presente en la sangre.
- ✓ Lipoproteína de alta densidad: lipoproteínas que transportan el colesterol hasta el hígado.
- ✓ Lipoproteína de baja densidad: lipoproteína asociada a la enfermedad aterosclerótica
- ✓ Colesterol de No Lipoproteína de alta densidad: colesterol para control de dislipidemias.
- ✓ Presión arterial sistólica: valor máximo cuando el corazón se contrae
- ✓ Presión arterial diastólica: capacidad de los vasos sanguíneos de distenderse
- ✓ Índice de masa corporal: nivel de grasa en el cuerpo
- ✓ Fumador: identifica si el paciente fuma
- ✓ Alcohol: identifica si el paciente consume alcohol
- ✓ Tratamiento con estatinas: terapia farmacológica para disminuir el colesterol LDL (malo)
- ✓ Tratamiento antihipertensivo: terapia farmacológica para reducir la presión arterial
- ✓ Tratamiento con aspirina: terapia farmacológica en pacientes de prevención secundaria

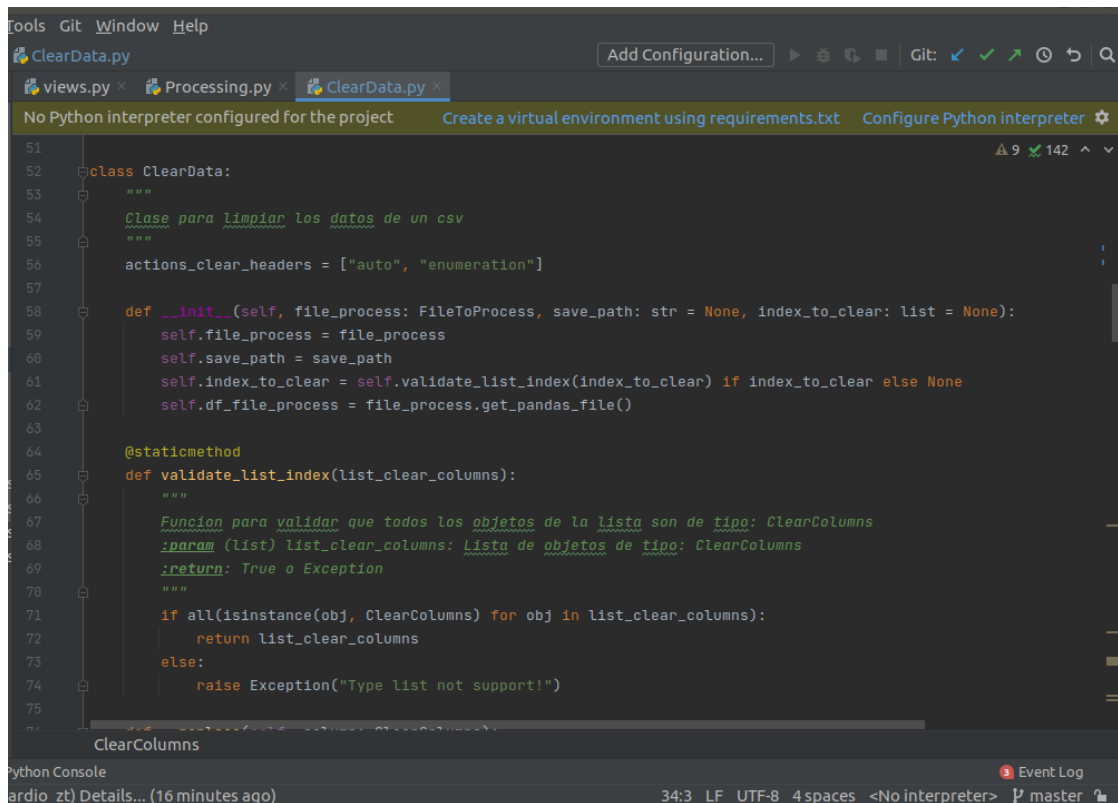
Consolidación de la información

Como resultado de la eliminación de variables que no se utilizaron para alimentar el sitio web a través del modelo de predicción y la adición de las variables complementarias, se obtuvo una base de datos idónea de la información de los pacientes a partir del año 2017, por ser la más relevante para este estudio (Ver Anexo 3).

6.2.3. Limpieza de los datos de las variables

Para esta actividad que correspondió a la limpieza de los datos de los pacientes seleccionados, esto es el caso específico de los pacientes con diabetes mellitus tipo 2 e

hipertensión arterial sistémica de la Consulta Externa del Hospital General Norte de Guayaquil “Ceibos”, donde se procedió de la siguiente manera. En primer lugar, luego de la recolección manual de los datos en el Hospital en Tablas de Excel (Ver Anexo 3) debido a que algunas variables no conservaban el mismo formato, se procedió a transformar la información de las mismas en valores cuantitativas para su correcto procesamiento en el modelo automático, para ello se utilizó la clase principal: ClearData. (Ver Figura 14.)



```
51
52 class ClearData:
53     """
54     Clase para limpiar los datos de un csv
55     """
56     actions_clear_headers = ["auto", "enumeration"]
57
58     def __init__(self, file_process: FileToProcess, save_path: str = None, index_to_clear: list = None):
59         self.file_process = file_process
60         self.save_path = save_path
61         self.index_to_clear = self.validate_list_index(index_to_clear) if index_to_clear else None
62         self.df_file_process = file_process.get_pandas_file()
63
64     @staticmethod
65     def validate_list_index(list_clear_columns):
66         """
67         Funcion para validar que todos los objetos de la lista son de tipo: ClearColumns
68         :param (list) list_clear_columns: Lista de objetos de tipo: ClearColumns
69         :return: True o Exception
70         """
71         if all(isinstance(obj, ClearColumns) for obj in list_clear_columns):
72             return list_clear_columns
73         else:
74             raise Exception("Type list not support!")
75
76     ClearColumns
```

Figura 14. Clase principal ClearData

Además aquí para la limpieza de la data, se procedió a eliminar los valores nulos, datos inconsistentes como caracteres no correspondientes al valor de la columna de la variable, quedando la información de cada columna de manera adecuada para su procesamiento en un solo tipo de formato de acuerdo a su cabecera correspondiente, antes de ingresarla al modelo de regresión lineal múltiple para la predicción del riesgo cardiovascular de los pacientes seleccionados, donde las variables utilizadas fueron: el sexo, la edad, la diabetes mellitus, el colesterol total, el colesterol de lipoproteína de alta densidad, la presión arterial sistólica, la variable fumador y tratamiento antihipertensivo. Todas estas en porcentajes de valores enteros, como se visualiza en la imagen. (Ver Figura 15.)

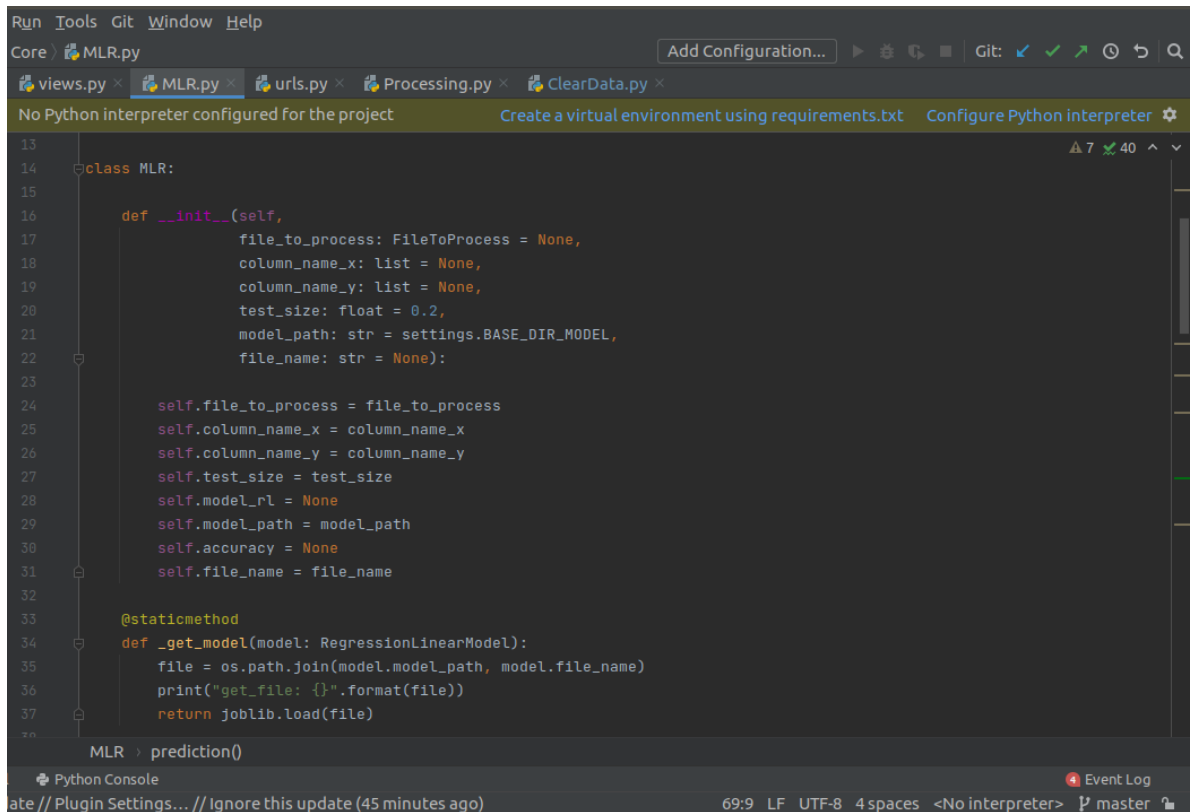
	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1	SEXO	EDAD	DMITZ	HTA	COL-T	TRIGLH	HDL-C	LDL-C	COL_NO	RAZA	PAS	PAD	IMC	FUMAI	ALCOH	CTO_EST	TTO_A	TTO_A	ENFER	IC	HOSPITA	MUERTE	CV		
2	F	61	1	1	275	235	46	183	229	1	121	62	30	NO	NO	SI	NO	NO	0	0	0	0			
3	F	41	1	1	259	139	48	192	211	1	116	69	29	NO	NO	NO	SI	NO	0	0	0	0			
4	M	44	1	1	186	194	49	99	137	1	114	62	37	NO	NO	SI	SI	NO	0	0	0	0			
5	F	37	1	0	153	129	47	107	106	1	144	83	39	NO	NO	NO	SI	NO	0	0	0	0			
6	F	63	1	1	360	317	65	128	295	1	127	90	35	NO	NO	NO	SI	NO	0	0	1	0			
7	M	82	0	1	147	119	44	68	103	1	100	80	17	NO	NO	SI	SI	NO	0	1	1	0			
8	F	60	0	0	152	214	44	101	108	1	115	65	32	NO	NO	SI	NO	SI	0	0	0	0			
9	F	49	0	1	195	114	55	77	140	1	154	89	40	NO	NO	NO	SI	NO	0	0	0	0			
10	F	72	0	1	271	132	47	198	224	1	135	80	34	NO	NO	NO	SI	NO	0	0	0	0			
11	M	39	1	0	206	139	47	135	159	1	124	83	30	NO	NO	SI	NO	NO	0	0	0	0			
12	F	81	0	1	158	368	40	129	118	1	93	50	21	NO	NO	SI	SI	NO	1	0	1	1			
13	M	38	1	0	246	153	48	149	198	1	120	70	30	NO	NO	NO	SI	NO	0	0	1	0			
14	M	48	1	1	302	280	54	110	248	1	130	70	28	NO	NO	NO	SI	NO	0	0	0	0			
15	M	50	1	0	214	284	40	129	174	1	100	80	33	NO	NO	SI	NO	NO	0	0	0	0			
16	F	63	1	1	179	205	40	106	139	1	120	80	34	NO	NO	NO	SI	NO	0	0	0	0			
17	F	31	1	0	234	189	44	168	190	1	123	71	29	NO	NO	SI	NO	NO	0	0	0	0			
18	F	60	1	1	167	99	55	102	112	1	120	60	28	NO	NO	NO	SI	NO	0	0	1	0			
19	F	65	1	0	200	205	50	142	150	1	109	55	33	NO	NO	SI	NO	NO	0	0	0	0			
20	F	84	0	1	193	134	54	112	139	1	110	62	25	NO	NO	SI	SI	NO	0	0	0	0			
21	M	52	1	1	141	184	43	92	98	1	120	80	30	NO	NO	NO	SI	SI	0	0	0	0			
22	M	57	1	1	187	482	45	112	142	1	155	86	32	NO	NO	SI	SI	NO	0	0	0	0			
23	M	62	0	1	132	108	48	73	84	1	124	68	30	NO	NO	SI	SI	NO	0	1	0	0			
24	M	78	0	1	145	459	44	190	101	1	107	77	22	NO	NO	NO	SI	SI	0	1	1	0			
25	M	61	1	0	156	280	43	90	113	1	110	73	31	NO	NO	NO	SI	NO	0	0	1	0			
26	F	60	1	0	175	136	60	106	115	1	100	67	27	NO	NO	NO	NO	NO	0	0	0	0			
27	F	63	1	1	188	331	46	87	142	1	109	72	29	NO	NO	SI	SI	NO	0	0	0	0			
28	F	40	1	0	208	268	52	154	156	1	104	84	36	NO	NO	NO	NO	NO	0	0	0	0			
29	F	23	1	0	160	128	55	80	105	1	117	65	35	NO	NO	NO	SI	NO	0	0	0	0			
30	F	59	1	0	180	90	65	97	115	1	121	60	31	NO	NO	NO	SI	NO	0	0	1	0			
31	M	48	1	0	233	195	42	172	191	1	118	85	22	NO	NO	SI	NO	NO	0	0	0	0			
32	M	46	1	0	167	178	45	122	122	1	120	60	23	NO	NO	SI	SI	NO	0	0	0	0			
33	F	43	1	0	212	111	62	128	150	1	104	62	22	NO	NO	SI	NO	NO	0	0	0	0			
34	M	44	1	0	127	110	53	60	84	1	110	60	24	NO	NO	NO	SI	NO	0	0	0	0			

Figura 15. Dataset transformado y limpio

Por lo tanto, los datos fueron transformados y limpiados correctamente con las funciones automáticas para así poder alimentar el modelo automático de regresión lineal de manera correcta al sistema web, utilizando las herramientas de Python: pandas, numpy y scikit-learn, para conseguir automatizar este proceso.

6.2.4. Diseño del modelo con aprendizaje automático.

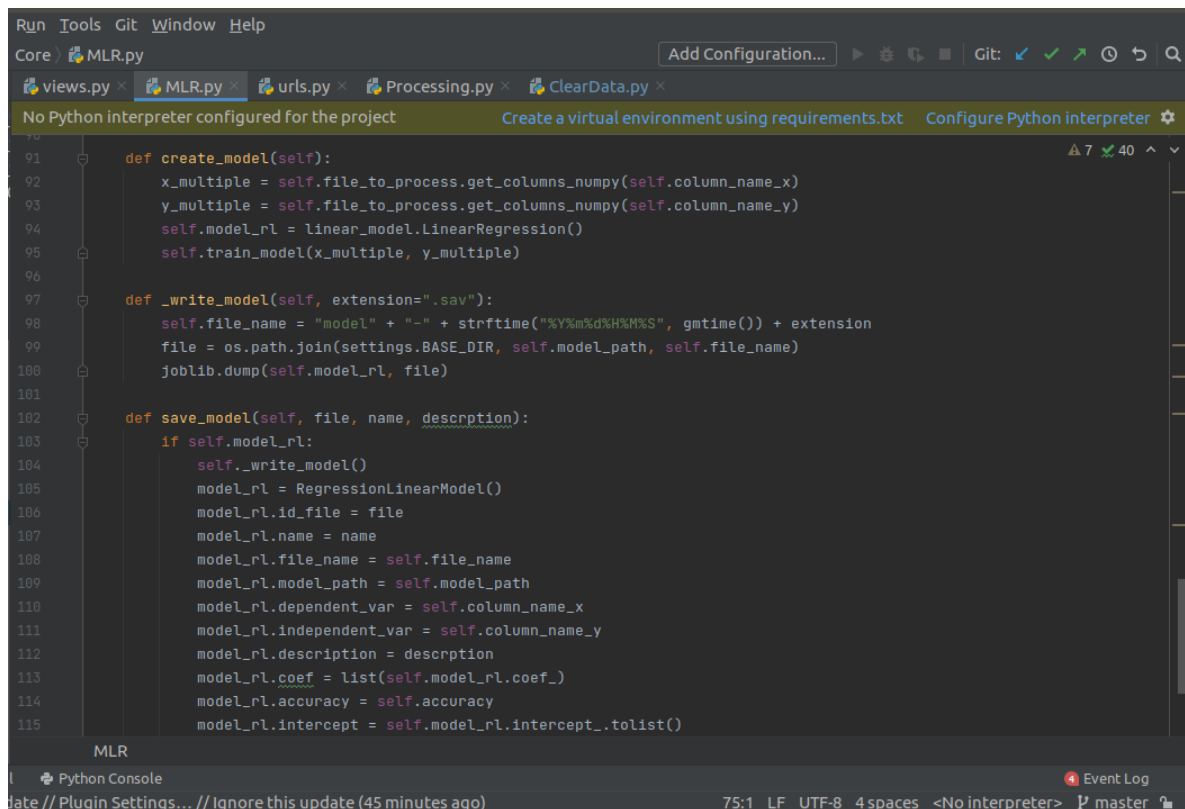
Para el diseño del modelo con aprendizaje automático (regresión lineal múltiple), se procedió de la siguiente manera. Primeramente, se creó el modelo de regresión lineal para el determinar el riesgo cardiovascular a través de una predicción en los pacientes seleccionados de su información correspondiente, donde su archivo principal fue MLR.py. (Ver Figura 16.)



```
13
14 class MLR:
15
16     def __init__(self,
17                 file_to_process: FileToProcess = None,
18                 column_name_x: list = None,
19                 column_name_y: list = None,
20                 test_size: float = 0.2,
21                 model_path: str = settings.BASE_DIR_MODEL,
22                 file_name: str = None):
23
24         self.file_to_process = file_to_process
25         self.column_name_x = column_name_x
26         self.column_name_y = column_name_y
27         self.test_size = test_size
28         self.model_r1 = None
29         self.model_path = model_path
30         self.accuracy = None
31         self.file_name = file_name
32
33     @staticmethod
34     def _get_model(model: RegressionLinearModel):
35         file = os.path.join(model.model_path, model.file_name)
36         print("get_file: {}".format(file))
37         return joblib.load(file)
38
```

Figura 16. Clase principal de modelo de regresión lineal múltiple

En la siguiente ilustración se detalla la utilización de la regresión lineal múltiple para el riesgo cardiovascular con sus funciones principales: `create_model`, `write_model`, `save_model`, entre otras, para predecir el valor en un porcentaje de 0 a 100%, donde su estado de respuesta del RCV se detalla en la sección 4.2.2 de la Revisión de Literatura (Ver Figura 17.)



```
Run Tools Git Window Help
Core > MLR.py Add Configuration...
views.py MLR.py urls.py Processing.py ClearData.py
No Python interpreter configured for the project Create a virtual environment using requirements.txt Configure Python interpreter
91 def create_model(self):
92     x_multiple = self.file_to_process.get_columns_numpy(self.column_name_x)
93     y_multiple = self.file_to_process.get_columns_numpy(self.column_name_y)
94     self.model_rl = linear_model.LinearRegression()
95     self.train_model(x_multiple, y_multiple)
96
97 def _write_model(self, extension=".sav"):
98     self.file_name = "model" + "-" + strftime("%Y%m%d%H%M%S", gmtime()) + extension
99     file = os.path.join(settings.BASE_DIR, self.model_path, self.file_name)
100     joblib.dump(self.model_rl, file)
101
102 def save_model(self, file, name, description):
103     if self.model_rl:
104         self._write_model()
105         model_rl = RegressionLinearModel()
106         model_rl.id_file = file
107         model_rl.name = name
108         model_rl.file_name = self.file_name
109         model_rl.model_path = self.model_path
110         model_rl.dependent_var = self.column_name_x
111         model_rl.independent_var = self.column_name_y
112         model_rl.description = description
113         model_rl.coef = list(self.model_rl.coef_)
114         model_rl.accuracy = self.accuracy
115         model_rl.intercept = self.model_rl.intercept_.tolist()
MLR
Python Console Event Log
date // Plugin Settings... // Ignore this update (45 minutes ago) 75:1 LF UTF-8 4 spaces <No interpreter> master
```

Figura 17. Creación del modelo de regresión lineal múltiple para el RCV

Se creó el modelo matemático para determinar el perfil lipídico del paciente con diabetes mellitus tipo 2 e hipertensión arterial sistémica, de acuerdo a la información específica recopilada de: colesterol total, hombre o mujer, lipoproteína de baja densidad, el colesterol de no lipoproteína de baja densidad y el valor de los triglicéridos.

6.3. FASE TRES: Predicción y Validación del Algoritmo Utilizado

Para la realización de esta fase tres, referente a la predicción y validación del algoritmo de aprendizaje automático de la regresión lineal múltiple utilizado en esta patología médica, se procedió de la siguiente manera:

6.3.1. Proceso

La propuesta del diseño del aplicativo web para determinar la relación entre el perfil lipídico y el riesgo cardiovascular en pacientes con enfermedades preexistentes como son la diabetes mellitus tipo 2 e hipertensión arterial sistémica, se compuso de:

- ✓ Análisis de variables de proyectos de aprendizaje automático para salud humana: aquí se realizó un análisis detallado de toda la información que podrían influir en el éxito del proyecto de predicción. Esta actividad fue parte de la primera etapa del flujo de trabajo.
- ✓ Flujo de Trabajo para el diseño del modelo: este flujo se realizó luego del análisis documental de mejores prácticas y procesos para diseñar un modelo de predicción.
- ✓ Selección del algoritmo: la selección se lo realizó de acuerdo al tipo de investigación (algoritmos de aprendizaje automático supervisado para predicción de enfermedades cardiovasculares), para el resultado que se requería obtener con sus consecuentes ventajas.

6.3.1.1. Planteamiento de la solución

En esta etapa se definen tres variables para la investigación: objetivo, origen de los datos, alcance y rendimiento.

- ✓ Objetivo: Predicción de éxito con el modelo de regresión lineal utilizado en el proyecto.
- ✓ Origen de los datos: Sistema Médico AS400 del Ministerio de Salud Pública.
- ✓ Alcance: Proyectos de machine learning entre 2017-2021
- ✓ Rendimiento del modelo: mínimo 70%.

6.3.1.2. Preparación de la data

Analizar variables y escoger la mayor cantidad posible

En esta etapa se llegó a destilar la información final de los pacientes que comprendían de lista inicial, donde se escogió con la parte médica las variables más idóneas de los datos para poderlas procesar.

Carga de la data

Se consolidó la información desde el Sistema Médico AS400 del Hospital Ceibos, el cual contenía los datos de los pacientes seleccionados en este estudio, para trabajarlos, a partir del período 2017, donde inició sus labores la entidad de salud.

Para poder asegurar que el modelo de predicción fuese confiable, se analizó cada una de las variables agregando y eliminando según correspondía su valor en la columna. Una vez determinadas las nuevas variables se procedió a registrar cada uno de los datos de los pacientes seleccionados. En lo que tuvo que ver a las variables categóricas que podían tener como valor de sí o no, esto para el caso de las variables: Sexo, DMT2, HTA, Fumador, Alcohol, Tto_Estatinas, Tto_Anti-HTA, Tto_Aspirina, Enfermedad_Adicional, IC, Hospitalización, Muerte_CV; para el caso especial de la variable raza sus valores fueron 0 para negra, 1 para mestiza y 2 para blanca. (Ver)

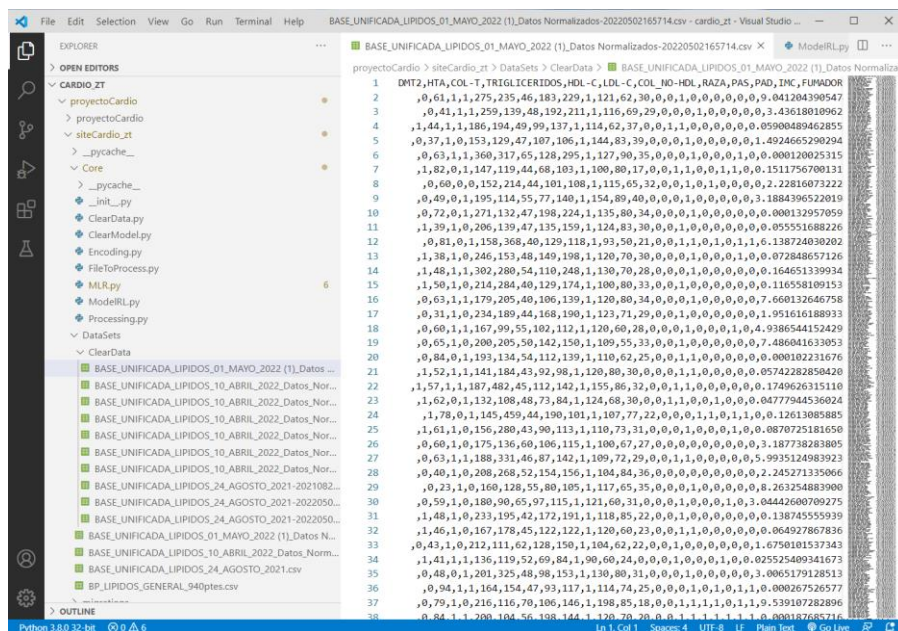


Figura 18 Dataset limpio del conjunto de pacientes seleccionados

Así de esta forma queda lista la información de los pacientes preparada para manipular en el modelo automático de la regresión lineal múltiple para el conjunto total de datos recopilados de un total de 5616 registros extraídos desde el programa AS400 del Hospital Ceibos.

Eliminación de datos null y variables no analizables

Una vez cargado el archivo, se procedió a verificar si existen campos que no contengan ningún valor y estén en blanco CSV (Ver Anexo 3). La herramienta presentó el resultado “False” que indica que no existen datos nulos, por lo que se procedió al siguiente paso, que fue el de eliminar las variables no analizables tales como códigos de diagnósticos presuntivos y códigos de diagnóstico definitivos y donde había inconsistencia en los datos.

6.3.1.3. Selección del algoritmo

En esta etapa se utilizó el algoritmo de aprendizaje supervisado para poder iniciar la predicción del modelo del riesgo cardiovascular y así obtener un valor porcentual del mismo como resultado. Se escogió el algoritmo de regresión lineal múltiple.

6.3.1.4. Entrenamiento del modelo

Cargar datos de entrenamiento al modelo. En esta etapa se alimentó el algoritmo de Regresión lineal múltiple con los datos destinados a entrenamiento, con un 80% para este modelo (Ver Anexo 3).

6.3.1.5. Validación del modelo

Cargar datos de evaluación al modelo. En esta etapa se alimentó al algoritmo de la regresión lineal múltiple con los datos destinados a evaluación en un 20% para que puedan ser predichas por el modelo automático.

Analizar métricas. En esta etapa se mostró los resultados del rendimiento y exactitud del algoritmo a la hora de predecir los datos de evaluación. Además, se determinó cuáles eran las variables con más peso para el éxito del modelo. (Ver Anexo 3).

6.3.2. Resultados

Los resultados presentados a continuación describen el rendimiento del modelo de predicción del algoritmo seleccionado. Además, se presentará el análisis de la regresión lineal de las variables con mayor peso para determinar los puntos más importantes tomados en cuenta a la hora de usar la información de los pacientes analizados.

En la figura 19, se visualiza la estructura del modelo predictivo con el algoritmo de regresión lineal múltiple alimentado por los datos de evaluación para el caso del paciente. Se contó con una variedad de variables para hacer al algoritmo más confiable a la hora de predecir el riesgo cardiovascular a diez años, verificando la precisión del modelo, con un porcentaje del 80% para entrenamiento y 20% para evaluación. Por tanto, el modelo de predicción es para la predicción del riesgo cardiovascular del paciente con enfermedad preestablecida.

The screenshot shows the 'CARDIO-ZT' web application interface. At the top, there is a navigation bar with 'Inicio', 'Servicios', 'Doctores', 'Contactos', and 'Salir'. Below this is a breadcrumb trail: 'Dashboard / Home / New Model'. The main content area is titled 'BASE_UNIFICADA_LIPIDOS_24_AGOSTO_2021-20210828181410'. It contains a form for creating a new model with the following fields:

- Nombre del modelo:** A text input field with the placeholder 'Nombre' and a note 'No mayor a 60 caracteres.'
- Porcentaje de entrenamiento:** A slider set to 20%, with a note 'No debe ser superior al 50%'.
- Descripción:** A large text area for entering a description.
- Variables dependientes:** A text input field with the placeholder 'VARIABLE.VARIABLE2' and a note 'Los nombres deben ser separados por comas.'
- Variables independientes:** A list of checkboxes for selecting independent variables: NOMBRE_PACIENTE, SEXO, EDAD, DMT2, HTA, COL-T, TRIGLICERIDOS, HDL-C, LDL-C, COL_NO-HDL, RAZA, PAS, PAD, IMC, FUMADOR, ALCOHOL, TTO_ESTATINAS, TTO_ANTI-HTA, TTO_ASPIRINA, ENFERMEDAD_ADICIONAL, IC, HOSPITALIZACION, MUERTE_CV, and CARDIO_RISK. A note below says 'Seleccione las columnas para crear el modelo.'

A blue 'Crear Modelo' button is located at the bottom left of the form area. At the bottom of the page, it says 'Created Aug. 28, 2021'.

Figura 19. Modelo predictivo para el riesgo cardiovascular

En la figura 20, se observa el resultado de la ejecución del código para poder visualizar en un nivel cuales son las variables con más peso a la hora de influir en el éxito de un proyecto. Así las variables de mayor peso son: COL-T, Sexo, HDL-C, No HDL, LDL-C, Triglicéridos.

The screenshot shows a web application interface for calculating a lipid profile. The header includes the logo 'CARDIO-ZT' and navigation links: 'Inicio', 'Servicios', 'Doctores', 'Contactos', and 'Salir'. A dark sidebar on the left contains menu items: 'Dashboard', 'Resultados', 'Listar Datos', 'Limpiar datos', and 'Perfil Lipídico'. The main content area has a breadcrumb trail 'Dashboard / Home / Perfil Lipídico' and a title 'Calcular Perfil Lipídico'. The form contains the following fields and values:

Variable	Valor
Colesterol total:	202
Sexo:	Mujer
HDL-C:	44
NO-HDL-C:	86
LDL-C:	56
Triglicéridos:	151

A blue 'Calcular' button is located below the input fields.

Figura 20: Variables para el cálculo del perfil lipídico

Mientras que en la figura 21, se observa los resultados de la ejecución del modelo predictivo con el algoritmo de Regresión Lineal Múltiple alimentado por los datos para el caso del perfil lipídico. Con este algoritmo se obtuvo una predicción de los datos recopilados, obteniendo un modelo confiable. Esto se debe a que es mayor el número de registros de pacientes para el algoritmo de regresión lineal múltiple requerido para entrenar los datos y así lograr el porcentaje de exactitud esperado. (Ver Figura 21)

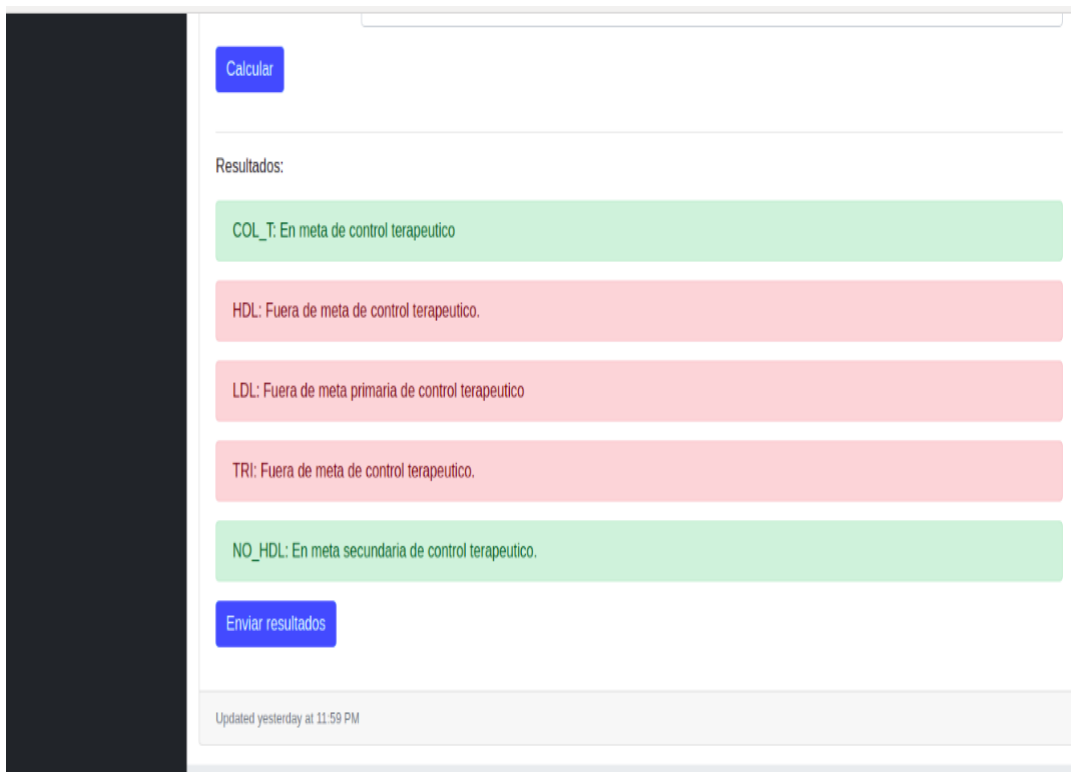


Figura 21. Variables para el cálculo del perfil lipídico

Al lograr un valor mayor al 70% de exactitud del modelo automático para la predicción del riesgo cardiovascular en los pacientes con diabetes mellitus tipo 2 e hipertensión arterial sistémica, con la utilización y selección de las variables más influyentes y con mayor peso a la hora del pronóstico médico realizado, se valida el modelo predictivo, teniendo un porcentaje de éxito favorable como se indica en la figura 31.

6.3.3. Pruebas

Para esta tarea final de este trabajo de investigación, se realizó las pruebas de aceptación con la colaboración de dos médicos, el Dr. Antonio Pozo García y Dr. Lenin Torres Y. para pronosticar mediante las variables de estudio recopiladas el riesgo cardiovascular y el perfil lipídico, los cuales ejecutándola y corrigiendo la herramienta informática dieron su aprobación y certificación del presente sitio web con aprendizaje automático, dando por concluida la etapa de pruebas con resultados positivos al comparar con el criterio de los expertos sobre la efectividad de la solución informática en el proyecto de investigación.

En el cuadro siguiente se muestran los criterios de validación para cada función

Tabla 14: Criterios de validación de usuario experto

Nro.	Título de Escenario	Evento	Resultados de Aceptación			
			No cumple nada	Cumple poco	Cumple algo	Cumple todo
1	Interfaz gráfica de usuario.	La interfaz del sitio web es amigable para el usuario.				X
2	Ingreso al Sistema.	El usuario administrador puede ingresar al sistema.				X
3	Gestión de ingreso de datos del paciente.	El usuario puede registrar los datos del paciente.				X
4	Visualización del resultado del paciente.	El usuario puede ver las estadísticas relacionadas a los resultados del riesgo cardiovascular y el perfil lipídico del paciente.				X
5	Gestión de los datos relacionados al paciente.	El usuario puede ver las estadísticas relacionadas a la base de datos de la investigación.				X

PRUEBAS UNITARIAS

En lo referente a las pruebas unitarias elaboradas, se las trabajó en el módulo del perfil lipídico con Unitest de Django, utilizando las siguientes variables de estudio: el colesterol total (Col-T), el sexo (Hombre y Mujer), la lipoproteína de alta densidad (Hdl-C), el

colesterol no hdl (No_Hdl-C), la lipoproteína de baja densidad (Ldl-C) y los triglicéridos. A continuación, en la figura, se visualiza el código empleado para el perfil lipídico en el aplicativo web.

```

1 # Create your tests here.
2 from django.test import TestCase
3 from siteCardio_zt.Core.Processing import Perfillipidico
4 from django.test import TestCase
5
6 class Perfillipidico(TestCase):
7     def test_if_get_col_t(self):
8         """Unittest of if get_col_t"""
9         pl = Perfillipidico(col_t=10).results()
10        self.assertEqual(pl[0]["status"], "success")
11
12    def test_else_get_col_t(self):
13        """Unittest of else get_col_t"""
14        pl = Perfillipidico(col_t=200).results()
15        self.assertEqual(pl[0]["status"], "danger")
16
17    def test_if_get_hdL_c(self):
18        """Unittest of if get_hdL_c"""
19        pl = Perfillipidico(men=True, hdl_c=60).results()
20        self.assertEqual(pl[1]["status"], "success")
21
22    def test_else_get_hdL_c(self):
23        """Unittest of else get_hdL_c"""
24        pl = Perfillipidico(men=False, hdl_c=30).results()
25        self.assertEqual(pl[1]["status"], "danger")
26
27    def test_if_no_hdL_c(self):
28        """Unittest of if no_hdL_c"""
29        pl = Perfillipidico(no_hdL_c=85).results()
30        self.assertEqual(pl[4]["status"], "success")
31
32    def test_elif_no_hdL_c(self):
33        """Unittest of elif no_hdL_c"""
34
35
36 Perfillipidico - test_elif_hdL_c()

```

Figura 22: Codificación de pruebas unitarias del perfil lipídico

Y a través de dichas pruebas se obtuvieron los resultados correctamente esperados de los niveles del perfil lipídico en dos categorías: en meta de control (nivel bueno) y fuera de meta de control (nivel malo) del paciente para todas las variables analizadas. Ver los resultados alcanzados en las pruebas realizadas en la Tabla 16.

Tabla 15: Pruebas Del Perfil Lipídico

Número de Prueba	Datos insertados para el perfil lipídico	Resultados
Prueba 1	400, H, 70, 128, 325, 330	Fuera de meta de control, en meta de control, fuera de meta de control, fuera de meta de control, en meta de control.

Prueba 2	325, M, 68, 90, 160, 257	Fuera de meta de control, fuera de meta de control, en meta de control, fuera de meta de control, fuera de meta de control.
Prueba 3	138, H, 50, 62, 142, 88	En meta de control, en meta de control, en meta de control, en meta de control.
Prueba 4	192, M, 47, 138, 176, 145	En meta de control, en meta de control, en meta de control, fuera de meta de control, en meta de control.
Prueba 5	537, H, 36, 117, 600, 501	Fuera meta de control, fuera de meta de control, fuera de meta de control, fuera de meta de control.
Prueba 6	175, M, 68, 192, 123, 107	En meta de control, en meta de control, fuera de meta de control, en meta de control, en meta de control.
Prueba 7	402, H, 56, 144, 252, 346	Fuera de meta de control, en meta de control, fuera de meta de control, fuera de meta de control, fuera de meta de control.

Así, se pudo obtener los resultados de las pruebas unitarias de los niveles del perfil lipídico del paciente analizado en las categorías que determinaron el estado favorable o no de acuerdo a los parámetros (variables) establecidos bajo el acompañamiento médico especializado en estos temas concernientes a las enfermedades cardiovasculares y así su

posterior toma de decisiones de acuerdo a los resultados obtenidos para el control y tratamiento a realizar. Se puede observar en la figura 23 el estado satisfactorio de las pruebas unitarias del perfil lipídico.

```
Archivo Editar Ver Buscar Terminal Ayuda
alex@alex-Lenovo-YOGA-C930-13IKB:~$ cd cardio_zt/
alex@alex-Lenovo-YOGA-C930-13IKB:~/cardio_zt$ cd proyectoCardio/
alex@alex-Lenovo-YOGA-C930-13IKB:~/cardio_zt/proyectoCardio$ source ~/virtualEnv/cardioEnv/bin/activate
(cardioEnv) alex@alex-Lenovo-YOGA-C930-13IKB:~/cardio_zt/proyectoCardio$ python manage.py test
Creating test database for alias 'default'...
System check identified no issues (0 silenced).
.....
-----
Ran 10 tests in 0.031s

OK
Destroying test database for alias 'default'...
(cardioEnv) alex@alex-Lenovo-YOGA-C930-13IKB:~/cardio_zt/proyectoCardio$
```

Figura 23: Estado de las pruebas unitarias

PRUEBAS DE RENDIMIENTO

Respecto a las pruebas de rendimiento del sitio web CardioZT se empleó la herramienta Apache JMeter, la cual permitió comprobar la respuesta del sitio, cabe mencionar que las pruebas se realizaron en un hosting por ende las respuestas de servicio son menos rápidas al de un servidor local.

Por medio de las pruebas de carga y estrés se logró determinar el funcionamiento del sitio web CardioZT, un total de 900 peticiones simultáneas fueron enviadas obteniendo los resultados que se presentan en las Figura 24.

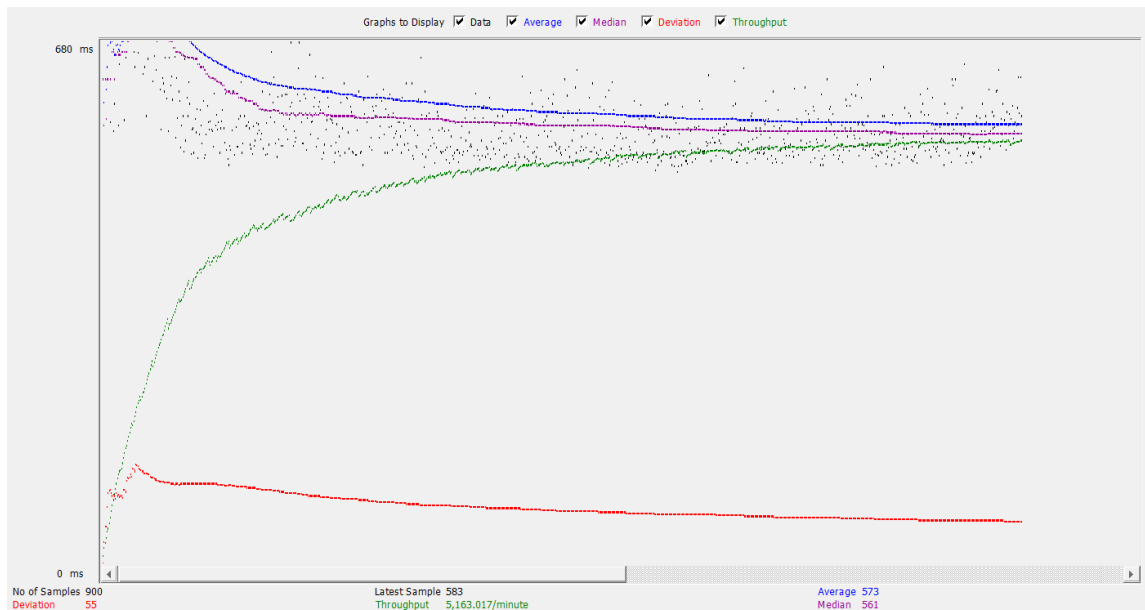


Figura 24: Prueba de rendimiento del sitio web CardioZT (Fuente: Apache JMeter)

De acuerdo a la figura que precede se presenta la Tabla 16 para una mejor comprensión, interpretando de esta forma la gráfica.

Tabla 16: Interpretación de la prueba de rendimiento

Campo	Valor	Interpretación
Muestras	900	Peticiones enviadas al servidor
Media	573	Tiempo promedio de respuesta en milisegundos
Min	510	Tiempo de respuesta más bajo obtenido
Max	1024	Tiempo de respuesta más alto obtenido
Error	0,00%	Porcentaje de Solicitudes con error
Rendimientos	2.1/sec	Tiempo transcurrido en realizar las 900 peticiones
Kb/seg	64.45	Rendimiento en kilobytes por segundo de la muestra

EXPERIMENTACIÓN

Introducción

Los datos del dataset unificado, fueron recopilados de manera manual de los pacientes a partir de los años 2017 al 2020, en un promedio de 8 horas de recolección de la información en los consultorios de los médicos especialistas en Medicina Interna del Hospital, con un total de 23 columnas y 5617 filas, en un formato csv.

Descripción de las variables seleccionadas del dataset unificado

La información de los pacientes recopilada bajo la dirección de la parte médica especialista fue de las siguientes variables de estudio, en la cual se da una descripción de cada columna utilizada en la presente investigación.

Tabla 17: Variables Del Conjunto Total De Datos Recopilados

Nombre de la Columna	Descripción	Rangos
HISTORIA CLINICA	Identificación única del paciente	
NOMBRE_PACIENTE	Nombre del paciente seleccionado	
SEXO	Género del paciente	Masculino (1), Femenino (0)

EDAD	Edad del paciente en años	20 a 98 años
DMT2	Diabetes Mellitus Tipo 2	Tiene (1), No tiene (0)
HTA	Hipertensión Arterial Sistémica	Tiene (1), No tiene (0)
COL-T	Colesterol Total	66 a 612
TRIGLICERIDOS	La fracción del colesterol, es un valor	18 a 970
HDL-C	Lipoproteínas de alta densidad. Se le llama colesterol "bueno"	40 a 70
LDL-C	Lipoproteínas de baja densidad. Se llama colesterol "malo"	40 a 200
RAZA	Etnia a la que pertenece: Negra (0), Mestiza (1), Blanca (2)	Negra (0), Mestiza (1), Blanca (2)
PAS	Presión Arterial Sistólica	90 a 220
PAD	Presión Arterial Diastólica	50 a 120
IMC	Índice de Masa Corporal	16 a 50
FUMADOR	Sí el paciente fuma (1), o no (0)	Fuma (1), No Fuma (0)
ALCOHOL	Sí el paciente ingiere alcohol (1), o no (0)	Toma alcohol (1), No toma alcohol (0)
TTO_ESTATINAS	Sí el paciente está recibiendo hipolipemiantes [píldoras]	Tiene tratamiento (1), No tiene tratamiento (0)
TTO_ANTI-HTA	Sí el paciente toma fármacos para la presión arterial	Tiene tratamiento (1), No tiene tratamiento (0)

TTO_ASPIRINA	Sí el paciente está recibiendo tratamiento con aspirina	Tiene tratamiento (1), No tiene tratamiento (0)
ENFERMEDAD_ADICIONAL	Sí el paciente tiene una enfermedad adicional	Tiene tratamiento (1), No tiene tratamiento (0)
IC	Paciente con Insuficiencia Cardíaca	Tiene tratamiento (1), No tiene tratamiento (0)
HOSPITALIZACIÓN	Sí el paciente ha sido hospitalizado	Tiene hospitalización (1), No tiene hospitalización (0)
MUERTE_CV	Sí el paciente ha fallecido	Paciente fallecido (1), Paciente no fallecido (0)

Desarrollo del Riesgo Cardiovascular

Modelo de Framingham: Este instrumento está asentado en el estudio del corazón de Framingham. Mide las posibilidades de sufrir un paro cardíaco en los próximos 10 años, en mayores de 20 años de edad.

Preproceso

Creamos el dataset con las variables necesarias para el pronóstico automático del riesgo cardiovascular del conjunto total de datos de los pacientes seleccionados para este estudio.

Nombre de las Variables:

- SEXO Sexo del paciente seleccionado.
- EDAD Edad del paciente.
- DMT2 Diabetes Mellitus Tipo 2.
- COL-T Colesterol Total.
- HDL-C Lipoproteína de alta densidad.

- PAS Presión arterial sistólica.
- FUMADOR Paciente es fumador o No.
- TTO_ANTI-HTA Tratamiento Antihipertensivo.

Análisis de la data y generación del modelo

En este apartado con la regresión lineal múltiple, se va a desarrollar un modelo para predecir el riesgo cardiovascular de los pacientes de acuerdo a: el colesterol total, la lipoproteína de alta densidad, el valor de la presión arterial sistólica, si tiene diabetes, es fumador; siendo estas las características que se tomaran en cuenta para diseñar nuestro modelo.

Para este análisis vamos a utilizar el dataset que desarrollamos manualmente en el Hospital “Ceibos”, correspondientes a los pacientes con diabetes mellitus tipo 2 e hipertensión arterial sistémica.

Primeramente, se importan las librerías que vamos a utilizar, importamos de sklearn el datasets, que es donde se guardan todos los conjuntos de datos que dispone esta librería y a su vez importamos de una vez linear_model.

```

1  from time import gmtime, strftime
2  from .FileToProcess import FileToProcess
3  from sklearn.linear_model import LinearRegression
4  from sklearn.model_selection import train_test_split
5  from sklearn import linear_model
6  from .ModelRL import ModelRL
7  from siteCardio_zt.models import PredictionsResult
8  from siteCardio_zt.models import RegressionLinearModel
9  from django.conf import settings
10 from sklearn.externals import joblib
11 import os

```

Figura 25: Importar librerías para el modelo de regresión lineal múltiple

Seguidamente iniciamos el método `__init__` de la clase Python, que tiene como objetivo inicializar los atributos del objeto que creamos dentro de la clase MLR, ya que es el primer método que se ejecuta cuando se crea un objeto.

```

class MLR:
    def __init__(self,
                 file_to_process: FileToProcess = None,
                 column_name_x: list = None,
                 column_name_y: list = None,
                 test_size: float = 0.2,
                 model_path: str = settings.BASE_DIR_MODEL,
                 file_name: str = None):

        self.file_to_process = file_to_process
        self.column_name_x = column_name_x
        self.column_name_y = column_name_y
        self.test_size = test_size
        self.model_path = None
        self.model_path = model_path
        self.accuracy = None
        self.file_name = file_name

```

Figura 26: Clase de Regresión lineal múltiple

Luego, con la función `_get_model` dentro de la clase `MLR`, se obtienen el modelo con la ruta y el nombre del archivo, es decir con los atributos `model.path` y `file.name`

```

33
34     @staticmethod
35     def _get_model(model: RegressionLinearModel):
36         file = os.path.join(model.model_path, model.file_name)
37         print("get_file: {}".format(file))
38         return joblib.load(file)

```

Figura 27: Obtención del modelo de regresión lineal

Seguidamente se inició el algoritmo de la regresión lineal múltiple con la función `start`. Obtenemos el contenido de la columna `X` con `get_columns_numpy` e imprimimos los valores de la variable `X_multiple`.

```

39
40     def start(self):
41         X_multiple = self.file_to_process.get_columns_numpy(self.column_name_x)
42         print(X_multiple)
43         print(type(X_multiple))

```

Figura 28: Función inicio de la regresión

Luego, se definen los datos correspondientes a las etiquetas en la variable `y_multiple`. Se separa los datos de entrenamiento y de prueba del algoritmo. Definimos el algoritmo a

utilizar, es decir la regresión lineal múltiple. Entrenamos el modelo con la función fit y realizamos una predicción con la función predict.

```

52 # Define los datos correspondientes a las etiquetas
53 y_multiple = self.file_to_process.get_columns_numpy(self.column_name_y)
54
55 # Separa los datos de "train" en entrenamiento y prueba para probar los algoritmos
56 X_train, X_test, y_train, y_test = train_test_split(X_multiple, y_multiple, test_size=self.test_size)
57 # Define el algoritmo a utilizar
58 self.modelo_rl = Linear_model.LinearRegression()
59 print("#####")
60 print(type(self.modelo_rl))
61
62 # Entrena el modelo
63 self.modelo_rl.fit(X_train, y_train)
64 # Realiza una predicción
65 Y_pred_multiple = self.modelo_rl.predict(X_test)

```

Figura 29: Definir los datos correspondientes a las etiquetas

Sabiendo toda esta información entonces procedemos a preparar los datos que vamos a utilizar para crear el modelo.

Para el caso del riesgo cardiovascular, se implementa un predictor de Regresión Lineal Múltiple, donde se tomará varios atributos del dataset para crear el modelo. Así solamente se tomará las columnas correspondientes al sexo correspondiente a la columna “SEXO”, la edad del paciente correspondiente a la columna “EDAD”, la columna de Diabetes Mellitus Tipo 2 correspondiente a el nombre “DMT2”, el Colesterol Total correspondiente a la columna “COL-T”, la lipoproteína de alta densidad correspondiente a la columna “HDL-C”, la presión arterial sistólica correspondiente a la columna “PAS”, fumador correspondiente a la columna “FUMADOR” y el tratamiento antihipertensivo correspondiente a la columna “TTO_ANTI-HTA”. Estas columnas vendrían a ser las variables regresoras o independientes correspondientes a la variable “X”.

```

57 print('PREDICCIÓN DEL MODELO DE REGRESIÓN LINEAL MÚLTIPLE')
58 print(Y_pred_multiple)
59 print('DATOS DEL MODELO REGRESIÓN LINEAL MÚLTIPLE')
60 print()
61 print('Valor de las pendientes o coeficientes "a":')
62 print(self.modelo_rl.coef_)
63 print('Valor de la intersección o coeficiente "b":')
64 print(self.modelo_rl.intercept_)
65 print('Precisión del modelo:')
66 self.accuracy = self.modelo_rl.score(X_train, y_train)
67 print(self.modelo_rl.score(X_train, y_train))

```

Figura 30: Estructura de la regresión lineal múltiple

Para el entrenamiento utilizamos la función train_model haciendo uso de la instrucción train_test_split, no sin antes importando la respectiva librería. Para la separación de los

datos, vamos a tomar un 20% de los mismos para utilizarlos como prueba una vez que hayamos obtenido el modelo.

```
76     def train_model(self, x_multiple, y_multiple, model=None):
77
78         # Separa los datos de "train" en entrenamiento y prueba para probar los algoritmos
79         X_train, X_test, y_train, y_test = train_test_split(x_multiple, y_multiple, test_size=self.test_size)
80         if model:
81             model.fit(X_train, y_train)
82             self.model_rl = model
83         else:
84             self.model_rl.fit(X_train, y_train)
85
```

Figura 31: Función de entrenamiento del modelo

Para establecer una comparativa de la predicción del método automático presentado en esta investigación versus un método tradicional del riesgo cardiovascular normalmente empleado por los especialistas médicos en estas patologías de enfermedades no transmisibles como son las enfermedades cardiovasculares, se realizó el siguiente procedimiento.

Primeramente, una comparativa de los resultados reales versus los resultados automáticos, donde se obtuvo un porcentaje de acierto de 85,23% de una muestra utilizada como se visualiza en la siguiente figura. (Ver Figura 29).

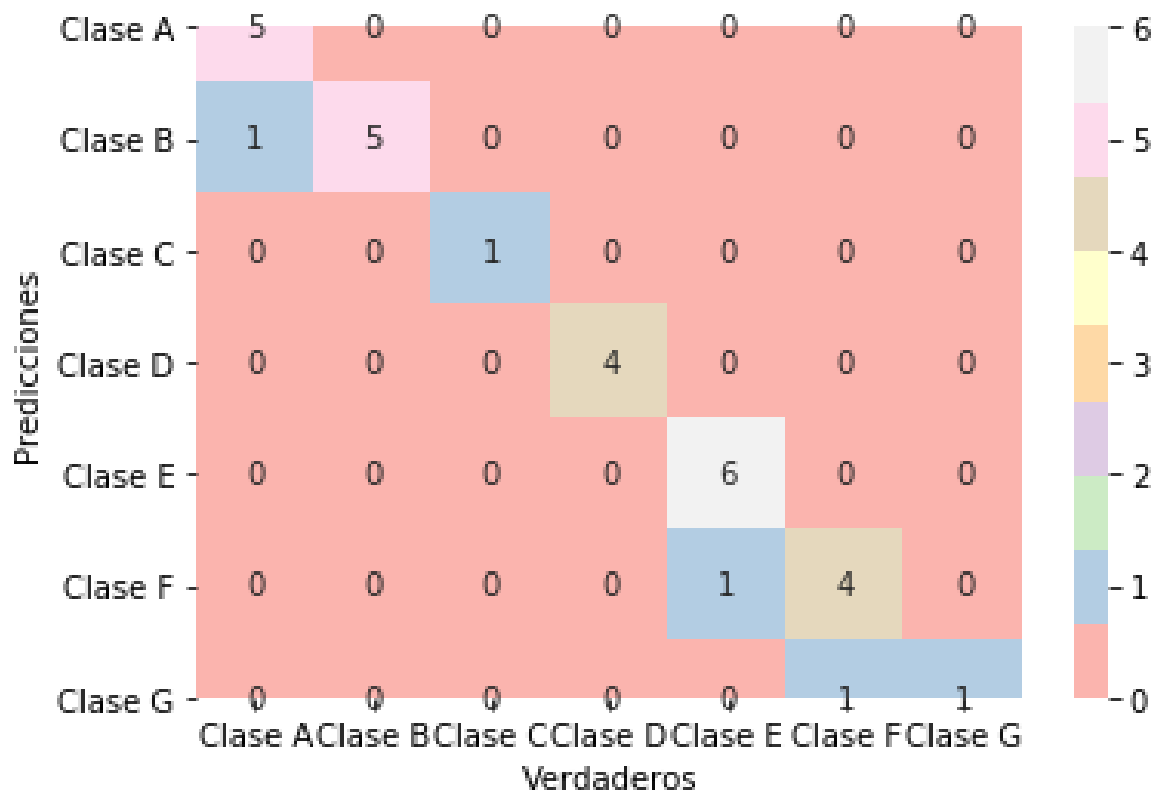


Figura 33. Matriz de confusión Nro. 1

Así, se puede observar en la figura 32, el adecuado funcionamiento del modelo de regresión lineal múltiple para los pacientes diabéticos tipo 2 e hipertensos en este trabajo investigativo, concerniente al perfil lipídico y su relacionamiento con el riesgo cardiovascular. En la figura 31 se muestra otra prueba efectividad del modelo con la matriz de confusión nro. 2 de los datos seleccionados para ver la funcionalidad del modelo con sus respectivos valores de datos reales versus datos automáticos con el uso del algoritmo de la regresión lineal múltiple

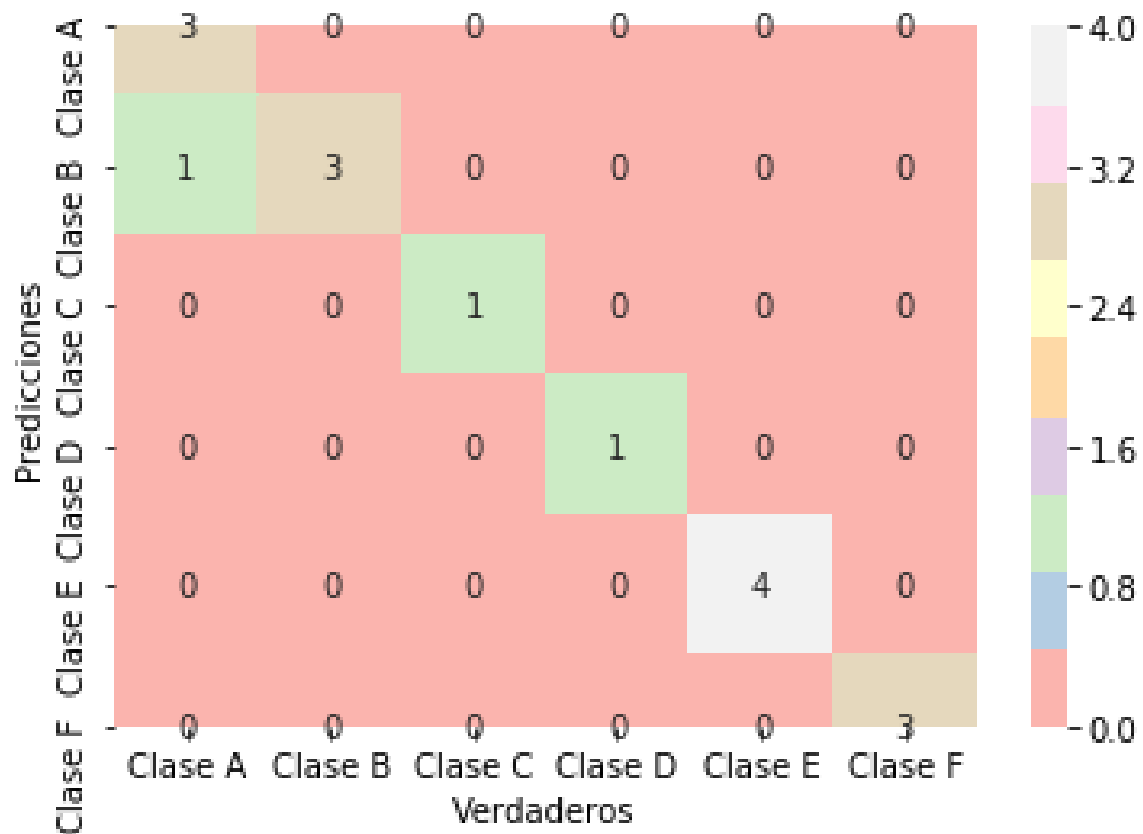


Figura 34: Matriz de Confusión Nro.2

Desarrollo del Perfil Lipídico

Para el desarrollo del Perfil Lipídico, bajo el criterio médico se seleccionaron las variables del dataset unificado: el colesterol total, la lipoproteína de alta densidad, el sexo, la lipoproteína de baja densidad y los triglicéridos, donde el resultado es en dos categorías: en meta de control y fuera de meta de control.

Preproceso

Crear el dataset con las variables necesarias

Nombre de las Variables:

- COL-T Colesterol Total.
- SEXO Sexo del paciente seleccionado.
- HDL-C Lipoproteína de alta densidad.
- LDL-C Lipoproteína de baja densidad.

- TRIGLICERIDOS Triglicéridos del paciente.

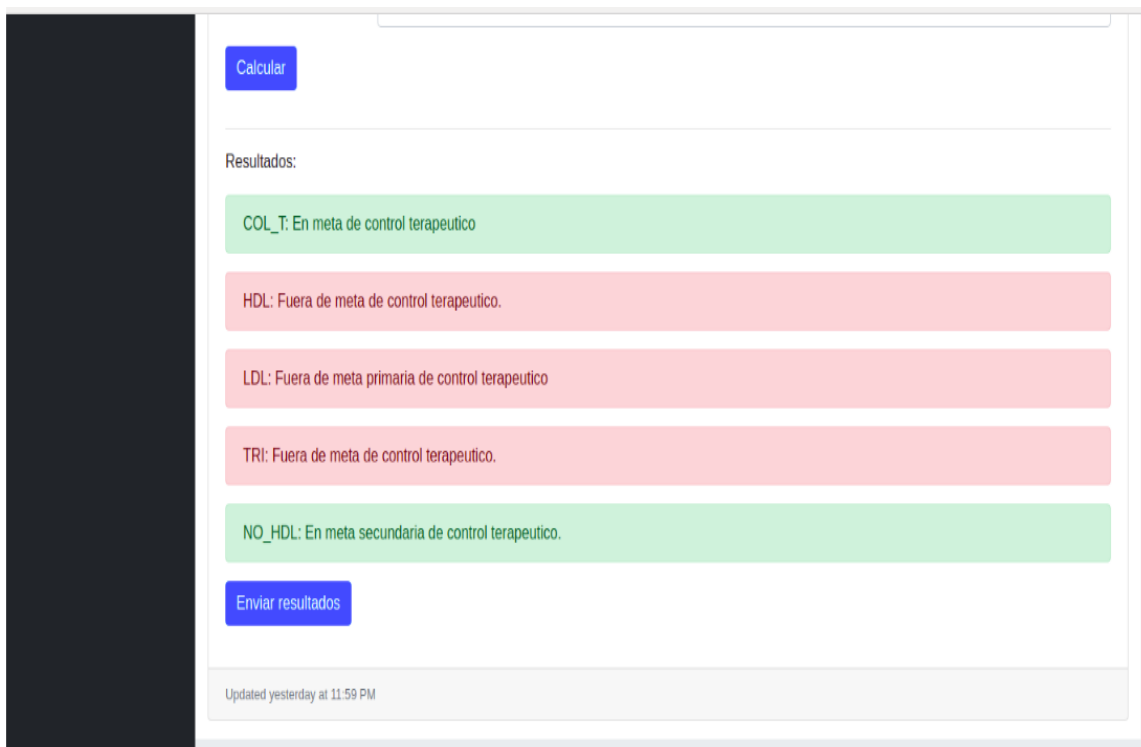


Figura 35. Niveles del Perfil lipídico del paciente.

Con los resultados obtenidos en la Fig. 34, en franjas color verde para el caso de un nivel bueno de componente del perfil lipídico, es decir en meta de control y en un color rojo cuando se encuentra fuera de meta de control. Finalmente, estos resultados son enviados a un correo electrónico donde se guardan para el respectivo análisis del médico especialista en Medicina Interna. La dirección del repositorio donde se aloja la aplicación web CardioZT está disponible en: <https://github.com/Roger281/cardiozt>

Proyección a futuro

Como proyección a futuro se podría mencionar la construcción de un aplicativo para el pronóstico de los eventos más rigurosos de las enfermedades cardiovasculares, como son: stroke, enfermedad renal crónica terminal con requerimiento de hemodiálisis, infarto agudo de miocardio, contando siempre con la apertura de las entidades de salud humana para el análisis de la información necesaria para estos estudios de relevancia social.

7. Discusión

Dentro del ámbito de la Medicina, la tecnología actual permite fomentar y propagar la información de esta área de conocimiento a través de herramientas como la Inteligencia Artificial y el Aprendizaje Automático, llegando a describir el desarrollo de una opción de solución experimental que especifique un modelo basado en el aprendizaje de máquina supervisado y manifieste los beneficios que se generan con dicha propuesta al grupo de pacientes analizado.

El presente trabajo de titulación comprendió una investigación de carácter exploratorio y retrospectivo, cuya finalidad fue definir un modelo predictivo para el pronóstico del riesgo cardiovascular y el cálculo del perfil lipídico de los pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos, para brindar al personal especialista de Medicina Interna la información pertinente del estado de salud de los pacientes seleccionados en esta investigación.

7.1. Revisar información bibliográfica de trabajos relacionados con este tipo de proyecto.

Para este apartado del proyecto de titulación el objetivo fue determinar la información relevante relacionada con la temática médica del perfil lipídico y el riesgo cardiovascular en pacientes diabéticos e hipertensos, concerniente a las enfermedades cardiovasculares. Para ello, dentro de los trabajos con mayor relevancia de acuerdo a los criterios de calidad y exclusión con el trabajo de titulación fueron, primeramente: “Desarrollo de un algoritmo para determinar el riesgo de muerte en pacientes dentro de una Unidad de Cuidado Intensivo utilizando Regresión Múltiple no Lineal” en el cual se trabajó con 3479 registros de pacientes de un total de 4000 bajo criterios de exclusión, donde la técnica con mayor desempeño para determinar el riesgo de muerte fue la máquina de vector de soporte, con puntajes entre 0.40 y 0.45 en un conjunto de validación completo. En la siguiente investigación: “Alfabetización en salud y control de la diabetes en pacientes de un Hospital Universitario de Argentina”, se tomaron 156 registros de pacientes seleccionados de un total de 349, en la cual se encontró una inadecuada Alfabetización en salud (AS) en un 60.3% de pacientes asociada con un mal control de la glucemia, a través de la técnica de la regresión lineal múltiple. Y en el trabajo investigativo: “Control metabólico y rendimiento escolar en niños con diabetes tipo 1”, se utilizaron 66 registros

de un total de 105 pacientes, en el cual se encontró una correlación con el número de controles diarios de glicemia capilar mediante el análisis de regresión lineal y se concluyó que el rendimiento escolar de pacientes con DM1 era inferior al de los niños sanos de la población general N-PG, debido al no control metabólico.

Así, se llegó a determinar que, de acuerdo al análisis de los trabajos relacionados, de todas las técnicas de aprendizaje automático supervisado, la regresión lineal múltiple era la más idónea para la consecución de este trabajo de titulación, sirviendo de esta manera para la predicción del riesgo cardiovascular y el cálculo del perfil lipídico en el conjunto total de pacientes recopilados.

Para mayor detalle revisar Resultados, Sección 6.1, y Ver sección 4.1

7.2. Recopilación y limpieza de datos para el perfil lipídico y el riesgo cardiovascular.

En esta fase concerniente a la recopilación y limpieza de los datos, primero se acudió personalmente al Hospital General del Norte de Guayaquil “Los Ceibos” en donde se realizó la entrevista pertinente a la parte médica especializada en el área de medicina interna con el objetivo de establecer los parámetros necesarios para trabajar en el sitio web y posterior a ello se comenzó a recolectar los datos necesarios de acuerdo a los criterios médicos para el inicio del trabajo investigativo, en el cual se recopiló 755 registros en tablas de Excel de los pacientes desde el año 2017 en primera instancia, luego con tres tablas adicionales de 4861 datos hasta completar el total de datos de 5616 registros finales para este estudio, recopilando la información de los niveles de colesterol, hábitos de salud y tratamientos a través de la historia clínica de cada paciente desde el programa AS400 del Ministerio de Salud Pública. Luego se tabularon los datos haciendo un filtrado de cada registro, respetando los criterios de inclusión del perfil lipídico y riesgo cardiovascular en el departamento de Consulta Externa en el piso 3 de la Torre 6 del Hospital “Los Ceibos”. Adicionalmente, se verificó que los datos estuvieran llenados correctamente en las columnas respectivas.

Los datos proporcionados se estructuraron en 16 columnas que constaban de los encabezados: fecha de atención, dependencia, nombre Medico, historia clínica, edad, sexo, raza, hipertensión arterial, presión arterial sistólica, presión arterial diastólica,

índice de masa corporal, diabetes mellitus, colesterol total, triglicéridos, lipoproteína de alta densidad, lipoproteína de baja densidad, colesterol no hdl, fumador, alcohol, tratamiento con estatinas, tratamiento anti hipertensión, tratamiento con aspirinas, código de diagnóstico presuntivo, diagnóstico presuntivo, código de diagnóstico definitivo y diagnóstico definitivo.

En la tabla 35, del Anexo 3, se muestra el archivo csv con 755 registros, además un archivo con la base general desde el año 2017. Estos datos fueron información real de pacientes hombres y mujeres, cuya edad es superior a 19 años, personas que ya tienen la enfermedad de diabetes mellitus e hipertensión arterial.

De igual forma en la tabla 36 del Anexo 3, se muestra el registro de la información de pacientes diabéticos y no diabéticos para: historia clínica, edad, sexo, índice de masa corporal, diabetes mellitus, colesterol total, triglicéridos, lipoproteína de alta densidad, lipoproteína de baja densidad.

Luego, en la tabla 37 del Anexo 3, se muestra la base general de los pacientes atendidos en los períodos 2017-2020, teniendo un registro mayor de 5000 pacientes, de las cuales finalmente se hizo uso de la técnica de regresión lineal múltiple para la predicción del riesgo cardiovascular y el perfil lipídico en pacientes que cumplen los criterios de inclusión.

Posterior a ello, en el sitio web se realizó una limpieza de datos con funciones automáticas, en donde la principal función principal fue `clear.data`, en la cual se limpiaron los datos de un archivo csv, se validaron que todos los objetos de la lista sean del mismo tipo, además utilizando otra función se reemplazó los valores vacíos de una columna, se creó una función para eliminar los valores vacíos, adicionalmente se limpiaron las cabeceras automáticamente, luego se enumeraron las cabeceras y se generó una función para establecer el valor de la variable, depurando así de esta manera los datos recopilados para su correcto ingreso con la regresión lineal múltiple en el modelo automático para el riesgo cardiovascular y el cálculo matemático del perfil lipídico del paciente, (Ver sección 6.2., subsección: 6.2.2 y 6.2.3.)

Se muestra a continuación en la Tabla 18, las categorías del riesgo cardiovascular para los pacientes seleccionados con las enfermedades preestablecidas de diabetes mellitus

tipo 2 e hipertensión arterial sistémica de la Consulta Externa del Hospital General Norte de Guayaquil Ceibos, que sirven de referencia médica para comprender la base científica de los parámetros establecidos y así obtener los resultados correctos del riesgo en el aplicativo.

Tabla 18: Estratificación del Riesgo Cardiovascular basado en M. Piepoli

CATEGORIAS DE RIESGO CARDIOVASCULAR	
Categoría	Pacientes con uno de lo siguiente:
Muy Alto RCV	<ul style="list-style-type: none"> ● ECV documentada. Clínica de ECV incluida antecedentes de IAM, SCA, revascularización coronaria y otros procedimientos de revascularización arterial, Stroke y AIT, aneurisma aórtico y EAP. ECV en imágenes incluye placa en la angiografía coronaria o ultrasonido carotídeo. No incluye un cierto aumento en los parámetros de imagen tales como el grosor de la íntima-media de la arteria carótida. ● DM con lesión de órganos diana tales como proteinuria o con un importante factor de riesgo como el tabaquismo o hipercolesterolemia o HTA marcada. ● ERC grave (TFG <30 ml/min/1,73 m²). ● Una puntuación calculada de SCORE \geq 10%.
Alto RCV	<ul style="list-style-type: none"> ● Factores de riesgo individuales marcadamente elevados, en particular el colesterol > 8 mmol/L (> 310mg/dl) (ej.: hipercolesterolemia familiar) o PA \geq180/110 mmHg. ● La mayoría de pacientes con DM (con la excepción de jóvenes con DM1 y sin factores de riesgo que pueden estar en riesgo bajo o moderado). ● ERC moderada (TFG 30-59 ml/min/1,73 m²).

	<ul style="list-style-type: none"> • Una puntuación calculada de SCORE $\geq 5\%$ Y $< 10\%$.
Moderado RCV	SCORE $\geq 1\%$ y $< 5\%$ a los 10 años. Muchos sujetos de mediana edad pertenecen a esta categoría.
Bajo RCV	SCORE $< 1\%$.
<p>ECV = enfermedad cardiovascular; SCA = síndrome coronario agudo; IAM = infarto agudo de miocardio; PA = presión arterial; ERC = enfermedad renal crónica; DM diabetes mellitus; TFG = tasa de filtración glomerular; EAP = enfermedad arterial periférica; SCORE = estimación sistemática de riesgo coronario; AIT = accidente isquémico transitorio.</p>	

Seguidamente, se visualiza en la Tabla 19, las metas de los factores de riesgo cardiovascular y niveles objetivo de los pacientes seleccionados con diabetes mellitus tipo 2 e hipertensión arterial sistémica del departamento de la Consulta Externa del Hospital General Norte de Guayaquil Ceibos, para determinar los resultados de los niveles del perfil lipídico del paciente que posee.

Tabla 19: Metas de Factores de Riesgo Cardiovascular basado en M. Piepoli

METAS DE FACTORES DE RIESGO CARDIOVASCULAR Y NIVELES OBJETIVO PARA IMPORTANTES FRCV	
Tabaco	No exposición en ninguna forma
Dieta	Baja en grasas saturadas con un enfoque en productos integrales, verduras, frutas y pescado.
Actividad Física (AF)	Al menos 150 minutos a la semana de AF moderada aeróbica (30 minutos durante 5 días/semana) o 75 minutos a la semana de

pacientes de Muy Alto RCV sin DM que pueden tolerar varios medicamentos para bajar la PA.

** No-HDL-C es un objetivo alternativo razonable y práctico, ya que no requiere ayuno. No-HDL-C con objetivos secundarios de <2.6, <3.3 y <3,8 mmol/l (<100, <130 y <145 mg/dl) se recomiendan para pacientes de muy alto, alto, moderado y bajo RCV, respectivamente.

*** Se dijo que los médicos de atención primaria podrían preferir un único objetivo general de LDL-C de 2,6 mmol/l (100 mg/dl). Si bien se acepta la simplicidad de este enfoque y que podría ser útil en algunos ajustes, hay un mejor apoyo científico para los tres objetivos coincidentes con el nivel de riesgo.

**** Esta es la recomendación general para los que están en muy alto riesgo. Debe tenerse en cuenta que la evidencia para pacientes con ERC es menos contundente.

7.3. Predicción y validación del algoritmo utilizado

Finalmente, todas estas partes previas de las fases anteriores como fueron la recopilación y limpieza de los datos se las integró en la aplicación web utilizando en el lenguaje de programación Python y el framework Django para el procesamiento de estos datos, utilizando la información recopilada de los pacientes en archivos csv, para poder emplearlos en el algoritmo de la regresión lineal múltiple y así de esta forma obtener los resultados de la predicción del riesgo cardiovascular en las métricas correspondientes con su respectiva etiqueta y para el caso del cálculo del perfil lipídico, en el cual los resultados se obtuvieron en las etiquetas determinadas en estados positivos o negativos del paciente (en meta o fuera de meta de control), de acuerdo a los expertos en medicina interna. (Ver en Experimentación de la sección 6.3, subsección 6.3.3.)

CARDIO-ZT Inicio Servicios > Doctores > Contactos Salir ↗

Dashboard / Resultados

Data Table Example

Primera tabla de predicciones...

Show 10 entries Search:

Características	Valor
_id_regresion_lm_cache	19, Modelo riesgo cardiovascular, date: 2021-08-28 18:16:56.724522+00:00
coef	[[1.31429352e-01 2.13845371e-03 1.55655381e-02 2.29964228e-04 -7.90322250e-04 1.30985033e-04 4.53468336e-03 -8.41260836e-03]]
details	Valor 6.199331%, se define como: Alto Riesgo
id_regresion_lm_id	19
intercept	[-0.13768092]
precision	0.08199331
prediction	6.20%
Características	Valor

Showing 1 to 7 of 7 entries Previous 1 Next

Figura 36. Riesgo cardiovascular con regresión lineal múltiple

En la Figura 33, se visualiza el resultado del modelo automático para la predicción del riesgo cardiovascular con su respectivo porcentaje y etiquetado, en el cual del conjunto total de datos recopilados se encontró un mayor porcentaje de registros para el caso del género femenino con un porcentaje 55.59%, y para el género masculino un valor de 44,41%.

De esta forma se llegó a resolver el problema de investigación, del cual se obtiene la relación existente entre el perfil lipídico del paciente diabético e hipertenso y su riesgo cardiovascular, para otorgar a los médicos especialistas en esta temática médica la información necesaria para la respectiva toma de decisiones en cuanto al tratamiento y control de dichos pacientes de la Consulta Externa del Hospital Ceibos de la ciudad de Guayaquil.

8. Conclusiones

- ✓ En esta investigación se llegó a determinar que la regresión lineal es una de las mejores técnicas de aprendizaje automático en base a los estudios científicos relacionados ya que ofrece los resultados más óptimos al momento de la predicción de factores asociados a la enfermedad cardiovascular, caso específico el riesgo cardiovascular, porque se integra adecuadamente al pronóstico médico.
- ✓ Con las técnicas de recopilación y limpieza de la información de la entrevista y la recopilación de información utilizadas en el proyecto se concluye que fueron las más adecuadas para su correcto ingreso al sistema web para el riesgo cardiovascular como para el cálculo del perfil lipídico, ya que la información pudo estar disponible de manera íntegra.
- ✓ En la aplicación con regresión lineal múltiple, se llegó a obtener un porcentaje de precisión de 79,27% referente al pronóstico del riesgo cardiovascular del paciente evaluado en la aplicación informática CardioZT.
- ✓ A través de la aplicación web se llegó a determinar que la relación existente entre el perfil lipídico y el riesgo cardiovascular es que a mayor alteración del perfil lipídico mayor será el riesgo cardiovascular en el paciente, de acuerdo a los resultados obtenidos en los módulos respectivos de funcionamiento.

9. Recomendaciones

- ✓ Se recomienda evaluar la información disponible en la web, bajo el criterio médico para su utilización en el análisis de enfermedades cardiovasculares, que se vinculen con el presente proyecto investigativo.
- ✓ Con el presente trabajo de titulación se pide a las Instituciones Hospitalarias, brindar las facilidades inherentes para la extracción y recopilación de los datos necesarios para hacer mejores modelos automáticos de las enfermedades cardiovasculares, caso especial niveles de colesterol.
- ✓ Se debería utilizar modelos computacionales por parte del personal médico en la detección de los factores preponderantes de las enfermedades con síndrome metabólico (cardiovascular) para dar la respuesta oportuna en función de los mejores diagnósticos de los pacientes.
- ✓ Para la construcción de un aplicativo informático más eficiente vinculado con la medicina humana se requeriría de la intervención de un grupo multidisciplinario con el suficiente contingente humano y logístico, para la prevención, control y seguimiento de patologías médicas.

10. Bibliografía

- [1] J. G. Corvalán, “Inteligencia artificial: retos, desafíos y oportunidades – Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia,” *Rev. Investig. Const.*, vol. 5, no. 1, p. 295, Jan. 2018, doi: 10.5380/rinc.v5i1.55334.
- [2] T. Boerma and C. D. Mathers, “The World Health Organization and global health estimates: Improving collaboration and capacity,” *BMC Med.*, vol. 13, no. 1, pp. 1–4, 2015, doi: 10.1186/s12916-015-0286-7.
- [3] D. Sisodia and D. S. Sisodia, “Prediction of Diabetes using Classification Algorithms,” *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 1578–1585, 2018, doi: 10.1016/j.procs.2018.05.122.
- [4] L. C. G. Martins, M. V. de O. Lopes, N. G. Guedes, A. P. de Menezes, O. de O. Farias, and N. A. dos Santos, “Árbol de clasificación para la evaluación del estilo de vida sedentario entre hipertensos,” *Investig. y Educ. en Enfermería*, vol. 34, no. 1, pp. 113–119, Feb. 2016, doi: 10.17533/udea.iee.v34n1a13.
- [5] R. A. Collazo *et al.*, “A COMPARATIVE STUDY BETWEEN ARTIFICIAL NEURAL NETWORK AND SUPPORT VECTOR MACHINE FOR ACUTE CORONARY SYNDROME PROGNOSIS,” *Pesqui. Operacional*, vol. 36, no. 2, pp. 321–343, Aug. 2016, doi: 10.1590/0101-7438.2016.036.02.0321.
- [6] M. Guerrero *et al.*, “Implementación de algoritmos de inteligencia artificial para la identificación de pacientes diabéticos utilizando los niveles de lípidos en sangre,” 2018.
- [7] O. D. Castrillón, W. Sarache, and E. Castaño, “Sistema Bayesiano para la Predicción de la Diabetes,” *Inf. tecnológica*, vol. 28, no. 6, pp. 161–168, 2017, doi: 10.4067/S0718-07642017000600017.
- [8] R. Guerrero-Pacheco and S. Galán-Cuevas Omar Sánchez-Armás Cappello, “Factores sociodemográficos y psicológicos asociados al autocuidado y la calidad de vida en adultos mexicanos con Diabetes Mellitus tipo 2 Sociodemographic and psychological factors associated with self-care and quality of life in Mexican adults with type,” 2017, doi: 10.14718/ACP.2017.20.2.8.

- [9] C. M. García García, J. D. Posada Aguilar, and J. Villanueva Padilla, “Desarrollo de un algoritmo para determinar el riesgo de muerte en pacientes dentro de una Unidad de Cuidado Intensivo utilizando Regresión Múltiple no Lineal,” *Prospectiva*, vol. 12, no. 2, p. 49, Nov. 2014, doi: 10.15665/rp.v12i2.288.
- [10] A. Sociedad Argentina de Investigación Clínica., M. E. Lanteri, C. Y. Cícero, A. Pérez, F. M. Puchulu, and R. Mejía, *ALFABETIZACIÓN EN SALUD Y CONTROL DE LA DIABETES EN PACIENTES DE UN HOSPITAL UNIVERSITARIO DE ARGENTINA*, vol. 77, no. 3. La Sociedad, 2017.
- [11] R. Román *et al.*, “Control metabólico y rendimiento escolar en niños con diabetes tipo 1,” *Rev Chil Pediatr*, vol. 88, no. 5, pp. 586–594, 2017, doi: 10.4067/S0370-41062017000500003.
- [12] F. P. Hernández Carrasco, “Desarrollo de una aplicación con redes neurales artificiales para el diagnóstico preventivo de la retinopatía diabética,” Jan. 2019.
- [13] L. Joyanes-Aguilar, N. J. Castaño, and J. H. Osorio, “Modelo de simulación y minería de datos para identificar y predecir cambios presupuestales en la atención de pacientes con hipertensión arterial Simulation and data mining model for identifying and prediction budget changes in the care of patients with hypertension,” *Rev. salud pública*, vol. 17, no. 5, pp. 789–800, 2015, doi: 10.15446/rsap.v17n5.3961.
- [14] F. Medina Rojas and C. Gomez, “Funcionalidades de la minería de datos,” *Ing. y Región*, vol. 12, no. 2, p. 31, Mar. 2015, doi: 10.25054/22161325.728.
- [15] L. Charris, C. Henriquez, S. Hernandez, L. Jimeno, O. Guillen, and S. Moreno, “Análisis comparativo de algoritmos de árboles de decisión en el procesamiento de datos biológicos,” Jun. 2018.
- [16] R. R. A. Santiago Leonardo Morales Cardoso, Mario Raúl Morales Morales, *Metodología para Procesos de Inteligencia de Negocios con mejoras en la extracción y transformación de fuentes de Datos*, vol. 4, no. 11 (2). MISC, 2015.
- [17] B. Ambale-Venkatesh *et al.*, “Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis,” *Circ. Res.*, vol. 121, no. 9, pp. 1092–1101, 2017, doi: 10.1161/CIRCRESAHA.117.311312.

- [18] J. P. R. Naivy Pujol Méndez, *Ciencia de datos: una revisión del estado del arte*, vol. 6, no. 3. Ichijinsha, 2017.
- [19] N. V. R. Pérez, M. L. Estrada, and N. N. R. Ramírez, “CLASIFICACIÓN SVM-DIFUSO,” *Pist. Educ.*, vol. 39, no. 125, Oct. 2017.
- [20] B. Benítez, C. Castro, R. A. Castañeda-Martínez, and A. Abaroa, “Predicción de Diagnóstico de Diabetes Mellitus utilizando Máquinas de Soporte Vectorial en Pacientes de Baja California,” *Memorias del Congr. Nac. Ing. Biomédica*, vol. 4, no. 1, pp. 415–418, Sep. 2017.
- [21] C. Lacave, A. I. Molina, M. A. Redondo, and M. Ortega, “Redes bayesianas para identificar perfiles de estudiante. Aplicación al estudio del abandono de las titulaciones de Informática en la Universidad de Castilla-La Mancha,” *Actas las XXII JENUI*, pp. 85–92, Jul. 2016.
- [22] E. M. M. Castro, “Bioestadística aplicada en investigación clínica: conceptos básicos,” *Rev. Médica Clínica Las Condes*, vol. 30, no. 1, pp. 50–65, Jan. 2019, doi: 10.1016/J.RMCLC.2018.12.002.
- [23] R. M. C. MJ Rodríguez-Jaume, “Análisis de regresión simple,” pp. 1–9, 2001.
- [24] F. Saldívar-González, F. D. Prieto-Martínez, and J. L. Medina-Franco, “Descubrimiento y desarrollo de fármacos: un enfoque computacional,” *Educ. Química*, vol. 28, no. 1, pp. 51–58, Jan. 2017, doi: 10.1016/J.EQ.2016.06.002.
- [25] E. Serna, *Desarrollo e innovación en energía*. 2017.
- [26] P. Ponce C, *Inteligencia artificial con aplicaciones a la ingeniería*. 2010.
- [27] L. V. Jaramillo and A. F. Antunes, “Detección de cambios en la cobertura vegetal mediante interpretación de imágenes Landsat por redes neuronales artificiales (RNA). Caso de estudio: Región Amazónica Ecuatoriana,” *Rev. Teledetección*, no. 51, p. 33, Jun. 2018, doi: 10.4995/raet.2018.8995.
- [28] D. A. Obando Velásquez, “Estudio y análisis de entornos comerciales mediante la evaluación, comparación y experimentación de algoritmos de minería de datos,” *Univ. Católica St. María - UCSM*, 2017.

- [29] P. Bambrick, W. S. Tan, R. Mulcahy, G. A. Pope, and J. Cooke, "Vascular risk assessment in older adults without a history of cardiovascular disease," *Exp. Gerontol.*, vol. 79, pp. 37–45, Jun. 2016, doi: 10.1016/j.exger.2016.03.002.
- [30] F. Mach *et al.*, "2019 ESC/EAS Guidelines for the management of dyslipidaemias: Lipid modification to reduce cardiovascular risk," *Eur. Heart J.*, vol. 41, no. 1, pp. 111–188, 2020, doi: 10.1093/eurheartj/ehz455.
- [31] K. Musunuru, "Atherogenic Dyslipidemia: Cardiovascular Risk and Dietary Intervention," *Lipids*, vol. 45, no. 10, pp. 907–914, Oct. 2010, doi: 10.1007/s11745-010-3408-1.
- [32] N. Townsend, M. Nichols, P. Scarborough, and M. Rayner, "Cardiovascular disease in Europe - Epidemiological update 2015," *Eur. Heart J.*, vol. 36, no. 40, pp. 2696–2705, 2015, doi: 10.1093/eurheartj/ehv428.
- [33] J. D. Berry *et al.*, "Lifetime Risks of Cardiovascular Disease," *N. Engl. J. Med.*, vol. 366, no. 4, pp. 321–329, Jan. 2012, doi: 10.1056/NEJMoa1012848.
- [34] J. T. Wilkins *et al.*, "Data Resource Profile: The Cardiovascular Disease Lifetime Risk Pooling Project," *Int. J. Epidemiol.*, vol. 44, no. 5, pp. 1557–1564, 2015, doi: 10.1093/ije/dyv150.
- [35] K. S. Reddy and A. Satija, "The Framingham Heart Study: Impact on the Prevention and Control of Cardiovascular Diseases in India," *Prog. Cardiovasc. Dis.*, vol. 53, no. 1, pp. 21–27, Jul. 2010, doi: 10.1016/j.pcad.2010.02.011.
- [36] S. Yusuf *et al.*, "Cardiovascular risk and events in 17 low-, middle-, and high-income countries," *N. Engl. J. Med.*, vol. 371, no. 9, pp. 818–827, 2014, doi: 10.1056/NEJMoa1311890.
- [37] F. Farzadfar *et al.*, "National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3·0 million participants," *Lancet*, vol. 377, no. 9765, pp. 578–586, Feb. 2011, doi: 10.1016/S0140-6736(10)62038-7.
- [38] M. T. Cooney *et al.*, "Re-evaluating the Rose approach: Comparative benefits of the population and high-risk preventive strategies," *Eur. J. Cardiovasc. Prev.*

- Rehabil.*, vol. 16, no. 5, pp. 541–549, 2009, doi: 10.1097/HJR.0b013e32832b38a1.
- [39] T. R. DAWBER, G. F. MEADORS, and F. E. MOORE, “Epidemiological approaches to heart disease: the Framingham Study.,” *Am. J. Public Health*, vol. 41, no. 3, pp. 279–281, 1951, doi: 10.2105/ajph.41.3.279.
- [40] G. L. Splansky *et al.*, “The Third Generation Cohort of the National Heart, Lung, and Blood Institute’s Framingham Heart Study: Design, recruitment, and initial examination,” *Am. J. Epidemiol.*, vol. 165, no. 11, pp. 1328–1335, 2007, doi: 10.1093/aje/kwm021.
- [41] M. J. Klag, D. E. Ford, and L. A. Mead, “The New England Journal of Medicine Downloaded from nejm.org at UNIVERSITY OF VICTORIA on August 5, 2014. For personal use only. No other uses without permission. Copyright © 2012 Massachusetts Medical Society. All rights reserved.,” 2012.
- [42] J. Stamler, R. Stamler, and J. D. Neaton, “Blood pressure, systolic and diastolic, and cardiovascular risks: US population data,” *Arch. Intern. Med.*, vol. 153, no. 5, pp. 598–615, 1993, doi: 10.1001/archinte.153.5.598.
- [43] J. T. DOYLE, T. R. DAWBER, W. B. KANNEL, A. S. HESLIN, and H. A. KAHN, “Cigarette smoking and coronary heart disease. Combined experience of the Albany and Framingham studies.,” *N. Engl. J. Med.*, vol. 266, pp. 796–801, 1962, doi: 10.1056/NEJM196204192661602.
- [44] C. S. Fox *et al.*, “Trends in cardiovascular complications of diabetes,” vol. 292, no. 20, pp. 2495–2499, 2016.
- [45] M. G. Goldschmid, E. Barrett-Connor, S. L. Edelstein, D. L. Wingard, B. A. Cohn, and W. H. Herman, “Dyslipidemia and ischemic heart disease mortality among men and women with diabetes,” *Circulation*, vol. 89, no. 3, pp. 991–997, 1994, doi: 10.1161/01.CIR.89.3.991.
- [46] W. B. Kannel, E. J. LeBauer, T. R. Dawber, and P. M. McNamara, “Relation of body weight to development of coronary heart disease. The Framingham study,” vol. XXXV, no. April, pp. 734–744, 1967.
- [47] M. Carrasco, W. Ocampo, L. Ulloa, and J. Azcona, “Metodología Híbrida De

- Desarrollo De Software Combinando Xp Y Scrum,” *Mikarimin. Rev. Científica Multidiscip.*, vol. 1, no. 1, pp. 17–28, 2019.
- [48] J. C. Salazar, Á. Tovar, J. C. Linares, A. Lozano, and L. Valbuena, “Scrum versus XP : similitudes y diferencias Scrum vs XP : Similarities and Differences,” *Tia*, vol. 6, no. 2, pp. 29–37, 2018.
- [49] Consejo de Educación Superior, “Reglamento De Regimen Academico,” *Regist. Of. Edición Espec. 854*, no. 111, pp. 1–51, 2019.
- [50] F. SÁENZ BLANCO ii FREDERICK GUTIÉRREZ SIERRA iii JULIÁN CAMILO RAMOS RIVERA iv CITACIÓN Sáenz-Blanco, “ESTABLISHMENT OF AGILE TEAMS FOR SOFTWARE DEVELOPMENT: REVIEW OF LITERATURE i CONFORMACIÓN DE EQUIPOS ÁGILES PARA EL DESARROLLO DE SOFTWARE: REVISIÓN DE LITERATURA FORMAÇÃO EQUIPAMENTO AGILE DESENVOLVIMENTO DE SOFTWARE,” 2017, doi: 10.15665/rde.v15i1+E1.1042.
- [51] M. E. Hossain, S. Uddin, and A. Khan, “Network analytics and machine learning for predictive risk modelling of cardiovascular disease in patients with type 2 diabetes,” *Expert Syst. Appl.*, vol. 164, no. August 2020, p. 113918, 2021, doi: 10.1016/j.eswa.2020.113918.
- [52] R. Caldera Vergara, “Estudio del framework de desarrollo web Django,” 2017.
- [53] D. A. Vera Paredes, L. C. Córdova Martínez, R. M. López Bermúdez, and ; Silvia Rosa Pacheco Mendoza, “Análisis de la metodología RUP en el desarrollo de software académico mediante la herramienta DJANGO,” *Recimundo*, vol. 3, no. 2, pp. 664–679, 2019, doi: 10.26820/recimundo/3.(2).abril.2019.964-979.
- [54] J. Zamorano Ruiz, “Comparación y análisis de métodos de clasificación con las bibliotecas scikit-learn y TensorFlow en Python,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [55] R. Quispe Llacotarimay, “Curso de Especialización en Machine Learning con Python,” 2020.
- [56] S. Moreno Fernández, “Herramienta de Reconocimiento de Imágenes en Python,”

2020.

- [57] J. R. Ruiz Sarmiento, F. Á. Moreno Dueñas, J. González Monroy, and A. J. Gonzalez-Jimenez, “Tutorial para el reconocimiento de objetos basado en características empleando herramientas Phyton,” Sep. 2018.
- [58] S. Arias and J. López, “Aplicación de la Inteligencia Artificial en la Bioinformática, avances, definiciones y herramientas.,” *UGCiencia*, pp. 159–171, 2016.
- [59] L. Sandoval Serrano, “Algoritmos de aprendizaje automático para análisis y predicción de datos,” pp. 36–40, 2018.
- [60] J. Sánchez Galán, “Machine Learning y sus Aplicaciones,” 2020.
- [61] M. N. Almeida Gachet, “MODELADO Y PREDICCIÓN DE LA CALIDAD DEL AIRE DE QUITO UTILIZANDO DATOS METEOROLÓGICOS Y DE CONTAMINACIÓN,” *Sustain.*, vol. 11, no. 1, pp. 1–14, 2019.
- [62] A. Carlos and J. Bahena, “Arquitectura para el despliegue tridimensional en dispositivos móviles de datos generados por tomógrafos .,” 2016.
- [63] J. L. Navarro Rey, “Desarrollo de una aplicación web con Django para evaluación automática de código Python,” 2017.
- [64] R. V. Baños, M. T. Fonseca, and M. R. Álvarez, “Análisis de regresión lineal múltiple con SPSS: un ejemplo práctico.,” *REIRE Rev. d Innovaci i Recer. en Educ.*, vol. 12 (2), pp. 1–10, 2019, doi: 10.1344/reire2019.12.222704.
- [65] M. D. L. Ángeles and H. Herrera, “Análisis Y Diseño De Un Sistema Para Identificar Signos De Retinopatía Hipertensiva A Través De Imágenes De Retina , Aplicando La tecnología de Deep Learning,” 2018.
- [66] S. Mendrano, “Modelo de minería de datos usando machine learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016 .,” 2016.
- [67] J. J. Mac Choy and A. R. Pumajulca Salazar, “Modelo Tecnológico Para El Control

Y Monitoreo Del Cuidado Médico De Enfermedades Crónicas Usando Wearables De Bajo Costo,” *Univ. Peru. Ciencias Apl.*, 2018.

- [68] O. Adrián *et al.*, “CONSTRUCCIÓN DE UN MODELO DE PREDICCIÓN PARA APOYO AL DIAGNÓSTICO DE DIABETES,” vol. 40, no. 130, pp. 2105–2122, 2018.
- [69] M. E. Andía, C. Arrieta, and C. A. Sing Long, “Una guía conceptual para usar y entender Big Data en la investigación clínica,” *Rev. Médica Clínica Las Condes*, vol. 30, no. 1, pp. 83–94, 2019, doi: 10.1016/j.rmclc.2018.11.003.
- [70] J. Ma *et al.*, “Assessment of triglyceride and cholesterol in overweight people based on multiple linear regression and artificial intelligence model,” *Lipids Health Dis.*, vol. 16, no. 1, pp. 3–9, 2017, doi: 10.1186/s12944-017-0434-5.
- [71] E. Kongshaug, “Association of Enhanced Perfusion Pattern Score with Type 2 Diabetes Severity Using Machine Learning Pipeline: Pilot Study,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [72] F. C. Mena Díaz, G. Nazar, and S. Mendoza Parra, “Antecedentes de adherencia a tratamiento en pacientes hipertensos de un centro de salud chileno,” *Hacia la Promoción la Salud*, vol. 23, no. 2, pp. 67–78, 2017, doi: 10.17151/hpsal.2018.23.2.5.
- [73] K. Singh and H. M. Chan, “Association of blood polychlorinated biphenyls and cholesterol levels among Canadian Inuit,” *Environ. Res.*, vol. 160, no. June 2017, pp. 298–305, 2018, doi: 10.1016/j.envres.2017.10.010.
- [74] C. D. Moulton, J. C. Pickup, A. S. Rokakis, S. A. Amiel, K. Ismail, and D. Stahl, “The Prospective Association Between Inflammation and Depressive Symptoms in Type 2 Diabetes Stratified by Sex,” *Diabetes Care*, no. 4, p. dc190813, 2019, doi: 10.2337/dc19-0813.
- [75] V. Alcalá, “Desarrollo de Biomarcadores basados en determinantes socioeconómicos y demográficos para el diagnóstico de Enfermedades Multifactoriales no transmisibles por medio de Redes Neuronales profundas,” 2019.

- [76] J. Santos, “Métodos de aprendizaje automáticos para detección de anomalías,” 2019.
- [77] J. Castellanos, A. Moreno, C. Bouza, and J. Sautto, “Valoración de Riesgo Cardiovascular mediante Modelos de Clasificación,” vol. 40, no. 1, pp. 80–87, 2019.
- [78] R. Baracaldo, “Sistema e-health para capturar, transmitir y almacenar el estado de la presión arterial y otros datos relevantes, con el fin de calcular la probabilidad de padecer hipertensión arterial,” 2019.
- [79] D. Tene and F. Montese, *Diseño de un prototipo web interactivo basado en técnicas de machine learning y razonamiento probabilístico para pronosticar mediante los factores de riesgo la enfermedad de diabetes*. 2019.
- [80] A. Macías and C. Cevallos, “Modelo basado en aprendizaje de máquina supervisado, para el análisis de datos cardiovasculares y su aplicación en el diagnóstico y pronóstico médico mediante el uso de una página web dinámica.,” 2019.
- [81] P. Dorado, “Contribuciones de las técnicas machine learning a la cardiología. Predicción de reestenosis tras implante de stent coronario,” 2019.
- [82] K. Eckert, S. D. Montenegro, N. López Forastier, and G. J. Candia, “Extracción de información de evoluciones clínicas digitales mediante técnicas de machine learning,” *X Congr. Argentino Informática y Salud - JAIIO 48 (Salta, 2019)*, pp. 60–74, 2019.
- [83] J. Sanaguano and A. Caina, “Desarrollo de una aplicación móvil para el monitoreo de pacientes con hipertensión arterial utilizando la tecnología ESP32,” 2020.
- [84] S. Retamar, A. D. E. Battista, L. Ramos, J. P. Nu, F. Savoy, and L. D. E. Gracia, “Minería de datos para la detección de factores de influencia en el test APGAR,” vol. 1, no. 6, 2019.
- [85] G. Maldonado, “Diabetes Mellitus Tipo 2 y sus efectos sobre procesos de Cognición Social,” 2019.
- [86] A. Aranda, “Estudio de la relación del esfuerzo de corte con la presión en

aneurismas cerebrales y la predicción del riesgo de ruptura usando herramientas de inteligencia artificial basado en datos morfológicos, fluidodinámicos y estructurales,” no. February, pp. 1–9, 2019, doi: .1037//0033-2909.126.1.78.

- [87] M. Barajas, D. Ballesteros, R. García, J. Pineda, and A. Buelvas, “Desarrollo de una aplicación móvil (Appteroma) que permita estimar el riesgo de padecer Aterosclerosis utilizando algoritmos de Machine Learning,” *Lis Solut.*, 2019.
- [88] P. I. Dorado-Díaz, J. Sampedro-Gómez, V. Vicente-Palacios, and P. L. Sánchez, “Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí,” *Revista Espanola de Cardiologia*, vol. 72, no. 12, Ediciones Doyma, S.L., pp. 1065–1075, 01-Dec-2019.
- [89] J. P. Tincopa Flores, “Diseño y evaluación de un sistema de ayuda al diagnóstico para neuropatía diabética basado en la lectura de puntos de presión plantar y machine learning,” pp. 1–83, 2019.
- [90] M. S. Cooper and M. Armentano, “Modelos de aprendizaje profundo para la predicción de comportamiento sedentario futuro,” vol. 19, no. 2, pp. 102–115, 2019.
- [91] F. Torres, E. Torres, J. Chile, D. De los Reyes, and N. González, “Análisis neutrosófico para el diagnóstico de la hipertensión arterial a partir de un sistema experto basados en casos,” *Rev. Fac. Ing.*, vol. 10, no. 60, pp. 202–213, 2019.
- [92] Y. García Montero, “Red neuronal para diagnóstico de hipertensión arterial,” p. 93, 2018.
- [93] C. Casas-Tapia, R. V. Araujo-Castillo, L. Saavedra-Tafur, A. Bert-Dulanto, A. Piscoya, and A. Casas-Lucich, “Higher HOMA-IR Index Is Associated With Increased Excess Weight Loss in Patients With BMI \geq 35 kg/m² After Vertical Gastrectomy,” *Cirugía Española (English Ed.)*, vol. 98, no. 6, pp. 328–335, 2020, doi: 10.1016/j.cireng.2020.06.003.
- [94] Y. Aguirre, “Métodos de aprendizaje supervisado para la predicción de diabetes: una revisión sistemática de literatura,” pp. 1–303, 2019.
- [95] A. Lagos, “Determinación de cristatina c e identificación de la relación con la

creatinina sérica en el diagnóstico de insuficiencia renal en pacientes con diabetes mellitus tipo 2,” 2017.

- [96] M. Aghaaminiha, S. A. Ghanadian, E. Ahmadi, and A. M. Farnoud, “A machine learning approach to estimation of phase diagrams for three-component lipid mixtures,” *Biochim. Biophys. Acta - Biomembr.*, vol. 1862, no. 9, p. 183350, 2020, doi: 10.1016/j.bbamem.2020.183350.
- [97] L. Hu, B. Liu, and Y. Li, “Ranking sociodemographic, health behavior, prevention, and environmental factors in predicting neighborhood cardiovascular health: A Bayesian machine learning approach,” *Prev. Med. (Baltim.)*, vol. 141, p. 106240, Aug. 2020, doi: 10.1016/j.ypmed.2020.106240.
- [98] A. Jamthikar, D. Gupta, N. N. Khanna, L. Saba, J. R. Laird, and J. S. Suri, “Cardiovascular/stroke risk prevention: A new machine learning framework integrating carotid ultrasound image-based phenotypes and its harmonics with conventional risk factors,” *Indian Heart J.*, vol. 72, no. 4, pp. 258–264, 2020, doi: 10.1016/j.ihj.2020.06.004.
- [99] K. Naseer Qureshi, S. Din, G. Jeon, and F. Piccialli, “An accurate and dynamic predictive model for a smart M-Health system using machine learning,” *Inf. Sci. (Ny)*, vol. 538, pp. 486–502, 2020, doi: 10.1016/j.ins.2020.06.025.
- [100] J. Goecks, V. Jalili, L. M. Heiser, and J. W. Gray, “How Machine Learning Will Transform Biomedicine,” *Cell*, vol. 181, no. 1, pp. 92–101, 2020, doi: 10.1016/j.cell.2020.03.022.
- [101] G. Bianchetti, F. Di Giacinto, M. De Spirito, and G. Maulucci, “Machine-learning assisted confocal imaging of intracellular sites of triglycerides and cholesteryl esters formation and storage,” *Anal. Chim. Acta*, vol. 1121, 2020, doi: 10.1016/j.aca.2020.04.076.
- [102] C. El-Hajj and P. A. Kyriacou, “A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure,” *Biomed. Signal Process. Control*, vol. 58, p. 101870, 2020, doi: 10.1016/j.bspc.2020.101870.
- [103] G. Thambiraj, U. Gandhi, U. Mangalanathan, V. J. M. Jose, and M. Anand,

- “Investigation on the effect of Womersley number, ECG and PPG features for cuff less blood pressure estimation using machine learning,” *Biomed. Signal Process. Control*, vol. 60, p. 101942, 2020, doi: 10.1016/j.bspc.2020.101942.
- [104] A. S. Alghamdi, K. Polat, A. Alghoson, A. A. Alshdadi, and A. A. Abd El-Latif, “A novel blood pressure estimation method based on the classification of oscillometric waveforms using machine-learning methods,” *Appl. Acoust.*, vol. 164, p. 107279, 2020, doi: 10.1016/j.apacoust.2020.107279.
- [105] G. Kissas, Y. Yang, E. Hwuang, W. R. Witschey, J. A. Detre, and P. Perdikaris, “Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks,” *Comput. Methods Appl. Mech. Eng.*, vol. 358, p. 112623, 2020, doi: 10.1016/j.cma.2019.112623.
- [106] Z. Gao *et al.*, “Learning physical properties in complex visual scenes: An intelligent machine for perceiving blood flow dynamics from static CT angiography imaging,” *Neural Networks*, vol. 123, pp. 82–93, 2020, doi: 10.1016/j.neunet.2019.11.017.
- [107] F. He *et al.*, “Novel Approach Combining Machine Learning and System Dynamics Modeling for Prediction of Risk of Secondary Cardiovascular Events and Effect of Interventions,” *J. Am. Coll. Cardiol.*, vol. 75, no. 11, p. 3545, 2020, doi: 10.1016/s0735-1097(20)34172-3.
- [108] P. Sengupta *et al.*, “Proposed Requirements for Cardiovascular Imaging Related Machine Learning Evaluation (PRIME) Checklist To cite this version : HAL Id : hal-03019705,” 2020.
- [109] O. Calvillo-Argüelles *et al.*, “Integrating the STOP-BANG Score and Clinical Data to Predict Cardiovascular Events After Infarction: A Machine Learning Study,” *Chest*, vol. 158, no. 4, pp. 1669–1679, 2020, doi: 10.1016/j.chest.2020.03.074.
- [110] R. Chen *et al.*, “Using machine learning to predict one-year cardiovascular events in patients with severe dilated cardiomyopathy,” *Eur. J. Radiol.*, vol. 117, no. June, pp. 178–183, 2019, doi: 10.1016/j.ejrad.2019.06.004.
- [111] A. Kilic, “Artificial Intelligence and Machine Learning in Cardiovascular Health

- Care,” *Ann. Thorac. Surg.*, vol. 109, no. 5, pp. 1323–1329, 2020, doi: 10.1016/j.athoracsur.2019.09.042.
- [112] M. Muthunayagam and K. Ganesan, “Cardiovascular Disorder Severity Detection Using Myocardial Anatomic Features Based Optimized Extreme Learning Machine Approach,” *Irbm*, vol. 1, no. Cvd, pp. 1–11, 2020, doi: 10.1016/j.irbm.2020.06.004.
- [113] S. Mezzatesta, C. Torino, P. De Meo, G. Fiumara, and A. Vilasi, “A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis,” *Comput. Methods Programs Biomed.*, vol. 177, pp. 9–15, 2019, doi: 10.1016/j.cmpb.2019.05.005.
- [114] A. Bansal *et al.*, “Machine Learning Techniques To Predict in-Hospital Cardiovascular Outcomes in Elderly Patients Presenting With Acute Myocardial Infarction,” *J. Am. Coll. Cardiol.*, vol. 75, no. 11, p. 3603, 2020, doi: 10.1016/s0735-1097(20)34230-3.
- [115] D. P. O’Regan, “Putting machine learning into motion: applications in cardiovascular imaging,” *Clin. Radiol.*, vol. 75, no. 1, pp. 33–37, 2020, doi: 10.1016/j.crad.2019.04.008.
- [116] D. Amaratunga, J. Cabrera, D. Sargsyan, J. B. Kostis, S. Zinonos, and W. J. Kostis, “Uses and opportunities for machine learning in hypertension research,” *Int. J. Cardiol. Hypertens.*, vol. 5, no. March, p. 100027, 2020, doi: 10.1016/j.ijchy.2020.100027.
- [117] O. Metsker, S. Sikorsky, A. Yakovlev, and S. Kovalchuk, “Dynamic mortality prediction using machine learning techniques for acute cardiovascular cases,” *Procedia Comput. Sci.*, vol. 136, pp. 351–358, 2018, doi: 10.1016/j.procs.2018.08.279.
- [118] C. J. Zack *et al.*, “Leveraging Machine Learning Techniques to Forecast Patient Prognosis After Percutaneous Coronary Intervention,” *JACC Cardiovasc. Interv.*, vol. 12, no. 14, pp. 1304–1311, 2019, doi: 10.1016/j.jcin.2019.02.035.
- [119] M. W. El Hechi, S. A. Nour Eddine, L. R. Maurer, and H. M. A. Kaafarani, “Leveraging interpretable machine learning algorithms to predict postoperative

- patient outcomes on mobile devices,” *Surg. (United States)*, vol. 169, no. 4, pp. 750–754, 2021, doi: 10.1016/j.surg.2020.06.049.
- [120] U. Benedetto *et al.*, “Machine learning improves mortality risk prediction after cardiac surgery: Systematic review and meta-analysis,” *J. Thorac. Cardiovasc. Surg.*, 2020, doi: 10.1016/j.jtcvs.2020.07.105.
- [121] F. Commandeur *et al.*, “Machine learning to predict the long-term risk of myocardial infarction and cardiac death based on clinical risk, coronary calcium, and epicardial adipose tissue: A prospective study,” *Cardiovasc. Res.*, vol. 116, no. 14, pp. 2216–2225, 2020, doi: 10.1093/cvr/cvz321.
- [122] A. M. Poss *et al.*, “Machine learning reveals serum sphingolipids as cholesterol-independent biomarkers of coronary artery disease,” *J. Clin. Invest.*, vol. 130, no. 3, pp. 1363–1376, 2020, doi: 10.1172/JCI131838.
- [123] P. Zalloua *et al.*, “Untargeted Mass Spectrometry Lipidomics identifies correlation between serum sphingomyelins and plasma cholesterol,” *Lipids Health Dis.*, vol. 18, no. 1, pp. 1–10, 2019, doi: 10.1186/s12944-018-0948-5.
- [124] M. J. Bom *et al.*, “Predictive value of targeted proteomics for coronary plaque morphology in patients with suspected coronary artery disease,” *EBioMedicine*, vol. 39, no. xxxx, pp. 109–117, 2019, doi: 10.1016/j.ebiom.2018.12.033.
- [125] M. M. Tomczyk and V. W. Dolinsky, “The cardiac lipidome in models of cardiovascular disease,” *Metabolites*, vol. 10, no. 6, pp. 1–19, 2020, doi: 10.3390/metabo10060254.
- [126] B. Ljubic *et al.*, “Predecir las complicaciones de la diabetes mellitus mediante algoritmos avanzados de aprendizaje automático,” *J. Am. Med. Informatics Assoc.*, vol. 0, no. 0, pp. 1–9, 2020, doi: 10.1093/jamia/ocaa120.
- [127] A. Aminian *et al.*, “Predecir el riesgo a 10 años de complicaciones en los órganos terminales de la diabetes tipo 2 con y sin cirugía metabólica: un enfoque de aprendizaje automático,” *Diabetes Care*, vol. 43, no. 4, pp. 852–859, 2020, doi: 10.2337/dc19-2057.
- [128] B. Pérez, “Explorando factores de riesgo de insuficiencia cardíaca a través del

aprendizaje automático,” pp. 2018–2019, 2019.

- [129] C. A. Jayashankar *et al.*, “Los niveles séricos de ácido úrico y colesterol unido a lipoproteínas de baja densidad son predictores independientes de enfermedad arterial coronaria en pacientes indios asiáticos con diabetes mellitus tipo 2,” *J. Nat. Sci. Biol. Med.*, vol. 7, no. 2, pp. 161–165, Jul. 2016, doi: 10.4103/0976-9668.184703.
- [130] H. Tada *et al.*, “Remnant-like particles and coronary artery disease in familial hypercholesterolemia,” *Clin. Chim. Acta*, vol. 482, no. March, pp. 120–123, 2018, doi: 10.1016/j.cca.2018.04.007.

11. Anexos

Anexo 1: Entrevista realizada al especialista de Medicina Interna.



UNIVERSIDAD NACIONAL DE LOJA

FACULTAD DE LA ENERGÍA, LAS INDUSTRIAS Y LOS RECURSOS NATURALES NO RENOVABLES

Estimado Doctor (a) la presente entrevista tiene como objetivo recolectar información para la ejecución del Trabajo de Titulación denominado “**Aplicación web para determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la consulta externa del Hospital Ceibos**”, la información con que usted aportará es de vital importancia, ya que con ello se realizará el levantamiento de requerimientos para el desarrollo del software.

¿Por qué es importante determinar el riesgo cardiovascular?

Es importante porque el riesgo cardiovascular engloba un conjunto de variables cualitativas y cuantitativas que permiten estimar la probabilidad de lesiones graves y o muerte a mediano y largo plazo en el endotelio vascular de órganos tan nobles como el corazón, cerebro, riñones los cuales están estrechamente interconectados y cuyo deterioro afecta de manera global a la salud del individuo.

¿Qué es el perfil lipídico?

Perfil lipídico en medicina se conoce al conjunto formado por las diferentes fracciones de colesterol que cumplen diversas funciones fisiológicas y/o fisiopatológicas (aterosclerosis) en el organismo humano según estén dentro o fuera

de límites cuantificables establecidos para la normalidad, de entre los cuales destaca el LDL-C por ser el elemento más estudiado, analizado en la mayoría de laboratorios y que se ha demostrado que mayor daño produce al interior de los vasos sanguíneos.

¿Cuál es el índice de mortalidad en el Ecuador por enfermedades cardiovasculares?

En Ecuador las enfermedades cardiovasculares son la primera causa de muerte, en el 2019 alcanzó el 26,49% del total de defunciones, según datos registrados por el MSP.

¿Cómo se trata al paciente con este tipo de enfermedades?

Se realiza un manejo multidisciplinario con intervenciones no farmacológicas y farmacológicas que se encaminan a incidir sobre factores de riesgos modificables y no modificables, entre los primeros destacan conductas tendientes a cambios en el estilo de vida que implican el dejar de fumar, abandonar el exceso de alcohol, evitar el sedentarismo y hacer ejercicio regularmente, descansar adecuadamente, llevar una alimentación saludable equilibrada evitando el exceso de grasas saturadas, carbohidratos complejos, gaseosas, controlando los niveles cuantificables de presión arterial, de azúcar, de colesterol; entre los no modificables mencionaremos a la edad, el sexo.

¿Existe una interacción médico-paciente para el control de la enfermedad?

Debería existir, ya que este aspecto resulta fundamental en el manejo de este tipo de enfermedades que, si bien no tienen un tratamiento curativo, permiten controlar la enfermedad y enlentecer la progresión acelerada hacia desenlaces fatales; un personal sanitario capacitado y un paciente que confía en ese personal se convierten en una poderosa combinación para enfrentar este tipo de patologías muy prevalentes en nuestra sociedad.

¿Cuáles son los tratamientos que reciben los pacientes?

En general el manejo es no farmacológico y farmacológico; en el primero se tiene los cambios en el estilo de vida; en el segundo aspecto se utilizan medicamentos conocidos como anti hipertensivos, anti diabéticos, estatinas y afines,

antiagregantes plaquetarios, anticoagulantes cuyas dosis y formas de administración varían según los requerimientos individuales de los afectados.

¿A qué edad comienza la persona a padecer esta enfermedad?

A lo largo de los años se ha podido documentar mediante biopsias endoteliales que la lesión inicial comienza en edades tempranas de la vida con la formación de placas de colesterol antes de los 5 años de edad, cuya progresión se produce sin dar manifestaciones clínicas significativas hasta la edad adulta; y, que en general a partir de los 40 a 50 años de edad comienzan las primeras manifestaciones clínicas de la enfermedad sumado a la alteración de los parámetros de laboratorio.

¿Qué propone para llevar un mejor control del paciente?

Insistir en una medicina preventiva para fomentar en los hogares, las escuelas, colegios estilos de vida saludables que ralenticen el inicio y la progresión de las enfermedades cardiovasculares; inculcar en la población en general el acudir al personal sanitario a valoraciones periódicas, programadas para despistaje de alteraciones que puedan incidir en alteraciones tempranas cardiovasculares; y, en los pacientes con patología ya establecidas brindar el entorno adecuado para el manejo oportuno, multidisciplinario, siguiendo protocolos establecidos por las diferentes sociedades médicas ecuatorianas y extranjeras.

¿Cuáles son los parámetros para determinar que una persona está controlado su riesgo cardiovascular?

Estos incluyen los siguientes: índice de masa corporal (IMC) normal entre 18 – 24, presión arterial <140/90 mmHg (<135/85 mmHg en DM), glicemia en ayunas <100 mg/dl, Colesterol total < 200 mg/dl, HDL-C >50 mg/dl, LDL-C < 70 mg/dl en HTA (< 56 mg/dl en DM), triglicéridos < 150 mg/dl.

¿Se puede enfermar una persona sana de estas patologías médicas: Diabetes Mellitus Tipo 2 e Hipertensión Arterial Sistémica?

La respuesta es sí, la carga genética heredada por alguno de nuestros progenitores nos predispone en promedio a un 30% de probabilidad de ello, y este valor se incrementa de acuerdo al entorno en el cual nos desenvolvemos y más aún si

tenemos niveles elevados de pobreza y falta de acceso a un sistema de salud eficiente.

¿Cuál es el presupuesto del Estado para este tipo de enfermedades?

Se desconoce la información respecto a este parámetro, ya que no existen datos fiables y de acceso público que informen el costo anual que representa para el Estado Ecuatoriano la atención de este tipo de patologías.

Anexo 2: Análisis bibliográfico de trabajos relacionados.

- ✓ ES01: En este proyecto: “Análisis y diseño de un sistema para identificar signos de retinopatía hipertensiva a través de imágenes de retina, aplicando la tecnología de deep learning”, se propone un prototipo de inteligencia artificial a través de la creación de un algoritmo para la limitación tecnológica en el área de oftalmología, que es la detección manual de la retinopatía hipertensiva. Su propósito es predecir por imágenes si una persona tiene la enfermedad o no. Para la creación de este sistema, se decidió por la metodología de creación de prototipos, ya que tiene etapas funcionales que sirven para crearlo. Trabajamos con la tecnología del aprendizaje supervisado a través del aprendizaje profundo y las redes neuronales convolucionales. Se utilizaron bases de datos como kaggle para obtener y estudiar imágenes retinianas. Debido a que las imágenes tenían diferentes tamaños y ruidos dentro de ellas, procedimos a preprocesar las retinografías. Después de procesar estas imágenes se pasan a una red neuronal de dos clases donde se estudian las características mínimas, que para una persona pueden tomar años de estudios. Se dio una solución a varias pruebas de sobre entrenamiento que se presentaron en las verificaciones. Al final, fue posible obtener un prototipo con una interfaz gráfica fácil de usar, que detecta si una persona tiene RHTA. Además de esto, el proyecto guarda las imágenes para hacer un seguimiento de los pacientes para una futura prevención de la enfermedad. Los recursos tecnológicos necesarios deben ser de alto rendimiento, recomendamos una máquina GPU NVIDIA. Este proyecto busca que las personas de Ecuador se adentren en el mundo de la inteligencia artificial y Big Data, ya que ahora son predicciones en imágenes de retina, pero mañana podrían ser fotografías de pulmones, cerebro o corazones, lo que ayuda a las personas a prevenir la enfermedad [65].
- ✓ ES02: En esta investigación: “Modelo de minería de datos usando machine learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016”, se desarrolló en la ciudad de Trujillo desde abril hasta diciembre del presente, centrándose en el problema de diagnósticos médicos y el uso de machine learning para contribuir en la mejora. La metodología que se usó

para desarrollar la investigación fue CRISP. El tipo de investigación de acuerdo al fin es aplicado, de acuerdo a la técnica de contrastación es Pre Experimental, la población es la ciudad de Trujillo, los instrumentos usados en la investigación son el cuestionario y el análisis documental.

En conclusión, se logró la mejora de los diagnósticos médicos a partir de un modelo de machine learning usando el algoritmo de árbol de decisión [66].

- ✓ ES03: En este trabajo: “Aplicación de la Inteligencia Artificial en la Bioinformática, avances, definiciones y herramientas”, la bioinformática, es el estudio de la información biológica a partir de la teoría de la información, la computación y las matemáticas; y, en un sentido más amplio, podría atribuírsele la función de las técnicas computacionales con el conocimiento, apreciación de los datos biológicos, almacenamiento, recuperación, manipulación y correlación de datos que proceden de distintas fuentes. El propósito encontrado en la actualidad es el de simplificar la complejidad y cantidad de datos analizados obtenidos a partir de la informática, tal es el caso, que se aplica en el diagnóstico clínico de enfermedades cardiovasculares.

Aquello surge debido a la problemática registrada en fenómenos tan complejos como lo son la genética, la farmacovigilancia, la probabilística. La característica común de todas ellas es que manejan un gran volumen de datos, con lo cual surge la necesidad de apoyarse en nuevas tecnologías informáticas para analizar esta avalancha de información; pero existen circunstancias en las que estas plataformas tecnológicas no están lo suficientemente desarrolladas para agilizar dichos procesos. Es aquí dónde surge el uso de herramientas, Técnicas, Frameworks, y metodologías de la inteligencia artificial para optimizar dichos procesos, con la consecuente reducción de tiempo y gasto informático que se producen en el manejo de los datos [58].

- ✓ ES04: En este trabajo: “Modelo tecnológico para el control y monitoreo del cuidado médico de enfermedades crónicas usando wearables de bajo costo”, MINSA es la institución comprometida con la prestación de salud a través del aseguramiento universal y un sistema de protección social. Sin embargo, el sistema de Salud Peruano presenta problemas de índole políticos, sociales y económicos que no permiten cubrir con la demanda de la población peruana, pues en el 2015 se registró que el 27.2% de personas no están afiliadas a un seguro de

salud en el Perú. Sin embargo, el rápido desarrollo de Sistemas Electrónicos que soportan la toma de decisiones Clínicas ha permitido la mejora de los servicios de salud en todo el mundo.

Por otro lado, las enfermedades crónicas se presentan como la mayor causa de mortalidad en el mundo debido al diagnóstico tardío y la inadecuada prevención. Por tal motivo, en este proyecto se realizó la implementación de un modelo tecnológico que sirve como sistema alternativo para el control y monitoreo de enfermedades crónicas mediante el uso de Wearables de bajo costo. Como consecuencia, la facilidad de construcción y la confiabilidad de su uso hacen viable que en el futuro la implementación del producto se realice a mayor escala, lo cual facilitará el acceso a servicios de salud más complejos a personas con limitaciones económicas y geográficas. Se realizó una investigación de los principales aportes de la Ingeniería de Sistemas y Electrónica en el campo del cuidado médico de pacientes con enfermedades no transmisibles para poder entender la relación entre los factores de riesgo fisiológicos de estas enfermedades y los dispositivos diseñados para controlarlas [67].

- ✓ ES05: En el presente estudio: “Construcción de un modelo de predicción para apoyo al diagnóstico de diabetes”, se procesaron variables como glucosa, insulina, presión arterial, edad, entre otras, para apoyo diagnóstico de Diabetes Mellitus; por medio de sistemas de aprendizaje automático y sistemas expertos con aprendizaje supervisado para la generación de árboles de decisión y análisis de resultados del algoritmo de predicción J48, con herramientas tales como BigML y Weka. Se llevó a cabo un procesado inicial de la información para determinar si el sujeto puede padecer la enfermedad. La variable principal fue la glucosa. La prueba se ejecutó con 768 registros de pacientes que han padecido o no la patología. Además, esto permitió al sistema tener una especificidad mayor al 70% para pacientes que no posee la enfermedad, y una sensibilidad del 63% para detección de la misma; los falsos positivos fueron 93 casos y los falsos negativos 108 casos, de un total de 500 pacientes analizados. Todo indica que el uso de la aplicación apoya satisfactoriamente a la presencia de la enfermedad [68].
- ✓ ES06: En el trabajo: “Una guía conceptual para usar y entender Big Data en la investigación clínica”, para estimar su impacto podemos utilizar como referencia

la revolución que constituyó en las últimas décadas la tecnificación de la medicina por sobre el arte de la medicina. El mejor ejemplo del potencial impacto de Big Data en medicina lo podemos encontrar en el estudio Framingham realizado en Massachusetts, Estados Unidos, cuyo seguimiento se inició en 1948 y actualmente continúa estudiando a la tercera generación de los participantes originales, a partir de este estudio se han sentado las bases fisiopatológicas y terapéuticas de muchas enfermedades cardiovasculares, nutricionales, entre otras. La investigación y práctica clínica se encuentran en una etapa de transición debido a nuevas metodologías para extraer información útil de un gran volumen de registros clínicos; y, que puede no ser evidente al utilizar los métodos tradicionales de análisis.

Como dato relevante tenemos que los métodos de análisis de datos utilizados presentan un incremento en la tasa de predicción de eventos cardiovasculares por sobre las guías médicas utilizadas hoy en día. Las características de Big Data son los seis elementos citados a continuación divididos en 2 grupos: volumen velocidad y variedad mismas que se refieren a la manipulación y administración de los datos; y, veracidad, variabilidad y valor que implican la relevancia de los datos en el contexto del análisis que se quiere realizar. Por otro lado, está la privacidad de la información de los pacientes y voluntarios involucrados. Existe una amplia adopción de nuevas técnicas en la práctica clínica en la que los profesionales de salud no sólo son generadores de datos, sino que, con el apoyo de las tecnologías emergentes, los datos, con su correcto análisis e interpretación, brinda un apoyo a los equipos de salud que permite tomar mejores decisiones médicas. La idea es crear grupos interdisciplinarios al interior de hospitales, clínicas y escuelas de medicina que permita a los profesionales de la salud familiarizarse con las nuevas técnicas de análisis desarrolladas, y esto a su vez permita a los profesionales que desarrollan fichas técnicas familiarizarse con las inquietudes y desafíos que enfrenta los profesionales clínicos teniendo como meta final el beneficio a los pacientes y usuarios de los sistemas de salud [69].

- ✓ ES07: En este estudio: “Evaluación de triglicéridos y colesterol en personas con sobrepeso basada en la regresión lineal múltiple y el modelo de inteligencia artificial”, se desea encontrar la relación de los triglicéridos (TG) y Colesterol (TC) con índices de la función hepática y renal. Donde se analizaron 364 personas

(varones y mujeres) con sobrepeso, todos fueron divididos en dos grupos: personas sanas y con sobrepeso; excluyendo personas con hepatopatía, angiocardopatía, nefrosis y enfermedades pulmonares, a través de distribución normal. Las relaciones de nivel de TG y TC con hígado, riñón y glucosa (GLU) se analizaron por correlación bivariada, donde los índices bioquímicos de hígado, riñón y glucosa se sometieron a un análisis de LRM, utilizando el software SPSS. Se usó la prueba Shapiro-Wilk y prueba de homogeneidad de las varianzas. El $p < 0,05$ se consideró estadísticamente significativo. Se utilizó análisis de mínimos cuadrados parciales para el modelo de regresión. Así, se encontraron 302 personas sanas y 273 con sobrepeso. En el exceso de peso hubo alta relación entre triglicéridos y el peso, altura, índice de masa corporal, glucosa y en colesterol se correlacionó con la edad. Por lo tanto, en el modelo de regresión lineal, la distribución residual es normal, siendo así fiable. [70].

- ✓ ES08: En este estudio: “Asociación de puntuación de patrón de perfusión mejorada con la gravedad de la diabetes tipo 2 usando la tubería de aprendizaje automático: estudio piloto”, dentro de la problemática de este trabajo cabe la duda razonable de la relación existente entre la severidad de la DMT2 y el daño neuronal que se traduce en posterior deterioro cognitivo. El objetivo es desarrollar una línea de investigación a través de aprendizaje automático para determinar el valor de discriminación del método comparando pacientes con DMT2 y casos control. Se realizó un estudio transversal y longitudinal prospectivo con un intervalo de 2 años. Población participante 131 individuos, de ellos 73 estaban divididos así: 41 con DMT2 y 32 sujetos sanos, de estos 42 individuos, 19 con DMT2 y 23 casos controles completaron los 2 años de seguimiento y se incluyeron en el análisis final. Se utilizó las siguientes variables: signos vitales, test de caminata, en sangre medición del hematocrito, glucosa, insulina, y hemoglobina glicosilada, se realizó test de aprendizaje y memoria: HVLTR (Hopkins Verbal Learning Test-Revised), y Mini-Mental State Examination (MMSE), medición de funciones ejecutivas: Verbal Fluency (VF), Trail Making (TM), Clock Drawing (CD), y medición de la atención: Digit Span (DS). Uso de protocolo con resonancia magnética nuclear con perfusión cerebral en 3D. Dentro de la evaluación: Aprendizaje automático para generar el patrón de red relacionado con DMT2 y los puntajes individuales asociados con el patrón.

Dentro de las pruebas estadísticas se hizo un análisis de regresión lineal con volumen de materia gris y años de educación como covariables. El método que aplica aprendizaje automático es superior al comparador con una mayor precisión de prueba 4%, área de la prueba bajo la curva 12%, valor predictivo positivo 7%, R de McFadden 24%. La puntuación del patrón de DMT2 de perfusión individual es un biomarcador de imagen de perfusión muy prometedor para rastrear la progresión de la enfermedad de pacientes individuales con DMT2 [71].

- ✓ ES09: En el artículo: “Antecedentes de adherencia a tratamiento en pacientes hipertensos de un centro de salud chileno”, se aborda un estudio cuantitativo de tipo correlacional para 141 pacientes hipertensos del programa de salud cardiovascular del centro sur de Chile. La falta de adherencia a tratamientos es un problema social, referentes a los servicios de salud, tipo de tratamiento e interacción médico-paciente, además de sus antecedentes sociodemográficos, antecedentes de salud, autoeficacia y estrés. La detección de la hipertensión futura es crucial a través de la predicción. Las variables de inclusión son ser mayor de 15 años y tener hipertensión en un período superior a 6 meses; las variables de exclusión fueron: presentar un deterioro cognitivo severo, no vivir en la comuna y presentar comorbilidades (diabetes y dislipidemia). El total de personas fueron 363, seleccionando 141 personas al azar. Para el análisis de datos se usó análisis de correlación, prueba Anova y regresión lineal múltiple, con un nivel de significación de 0,05 usando el programa SPSS 21 y R. Para la caracterización de los pacientes hipertensos se tomó una muestra de mujeres con un 75,9% y la media total de edad fue de 62 años. Para el comportamiento de las variables adherencia, autoeficacia y estrés, las puntuaciones son 30,7 para autoeficacia y 13,7 para estrés. Respecto a la predicción de adherencia al tratamiento, los hombres tienen mayor puntaje en la dimensión de seguimiento de indicaciones que las mujeres; en cambio en la subescala ingesta de medicamentos, las mujeres obtienen mayor puntaje que los hombres. En la subescala de seguimiento de dieta se encontró que las personas con más edad tienen mayor puntaje de adherencia en seguir la dieta que los jóvenes. Se llegó a determinar diferencias significativas en las escalas de seguimiento de indicaciones, consumo de medicamentos, manejo de estrés y en adherencia total respecto a si la persona es consciente de su cronicidad de la enfermedad [72].

- ✓ ES10: En este estudio: “Asociación de bifenilos policlorados en sangre y niveles de colesterol entre los Inuit Canadienses”, el propósito fue investigar si los bifenilos policlorados (PCB) en la sangre están asociados con el colesterol alto y los parámetros relacionados en los inuit canadienses, en treinta y tres comunidades costeras y tres comunidades del interior de la Región de Solución de Inuvialuit, territorio Nunavut y Nunatsiavut. Un total de 2595 inuit que eran mayores de 18 años de edad (68%) y 2191 con y sin colesterol alto; excluyendo mujeres embarazadas. Las covariables para la inclusión en los modelos fueron: edad, sexo, estado civil, educación, ingresos, alcohol, cigarrillo, ejercicio medido con la puntuación total equivalente metabólico (MET), índice de masa corporal (IMC), los niveles sanguíneos de metales pesados, ácidos grasos totales, ácidos grasos monoinsaturados (MUFA), ácidos grasos poliinsaturados (PUFA), grasas totales saturadas, ácidos grasos trans (TFA), omega-3 relación/omega-6, glucosa alta, presión sanguínea sistólica, presión arterial diastólica y la historia familiar. Se utilizaron modelos de regresión logística múltiple para examinar los bifenilos policlorados y el colesterol alto. Se realizaron modelos para triglicéridos, colesterol total, LDL-C y HDL-C respecto a la regresión lineal múltiple; para el colesterol alto se utilizó categorías de cuartil a base de lípidos, en el caso de los triglicéridos, colesterol total, LDL-C y HDL-C, los modelos se ejecutaron con niveles de cuartil de PCB de peso húmedo [73].
- ✓ ES11: En el artículo: “La asociación prospectiva entre la inflamación y los síntomas depresivos entre la Diabetes tipo 2 estratificada por sexo”, se pretende probar si la inflamación se relaciona con los síntomas depresivos en la diabetes y si es más fuerte en mujeres que hombres con diagnóstico reciente de diabetes tipo 2, en Londres. Así, los modelos se ajustaron por edad, origen étnico, índice de masa corporal, la presión sanguínea, colesterol, antidepresivos, medicamentos antiinflamatorios. El resultado a los 2 años fue síntomas depresivos bajo el Cuestionario de Salud Total del paciente (PHQ-9). Se usó una población de 3008 pacientes, donde se encontró que la elevada carga de inflamación se asoció con un mayor empeoramiento de los síntomas depresivos en el sexo femenino, así mismo el índice de masa corporal (IMC) aumentado, es otra razón de la vulnerabilidad femenina a la depresión. Lo que se sugiere es una terapia de género para reducir la carga de inflamación en las mujeres. Adicionalmente, se necesita

- un seguimiento más largo para confirmar la persistencia de la depresión inflamatoria en las mujeres, así como secuelas biomédicas asociadas [74].
- ✓ ES12: En este estudio: “Desarrollo de Biomarcadores basados en determinantes socioeconómicos y demográficos para el diagnóstico de enfermedades multifactoriales no transmisibles por medio de redes neuronales profundas”, se propone el uso de una red neuronal artificial que clasifique a las personas con las afecciones de diabetes y caries dental, usando 31 características de cada persona. El modelo fue evaluado por análisis estadístico a través de la precisión, función de pérdida, el área bajo la curva AUC, y la curva de características operativas receptoras ROC. Los resultados obtenidos tuvieron una precisión de 0,99, AUC de 0,99 para las curvas ROC [75].
 - ✓ ES13: En el artículo científico: “Implementación de algoritmos de inteligencia artificial para la identificación de pacientes diabéticos utilizando los niveles de lípidos en sangre”, se va a clasificar pacientes diabéticos con algoritmos de inteligencia artificial utilizando el colesterol total, HDL, LDH, y triglicéridos; del Centro Médico Siglo XXI, de 1019 registros. La metodología utilizada consta de dos etapas: entrenamiento y validación de los algoritmos; donde los modelos de regresión logística, árboles de decisión y máquinas de soporte vectorial, tienen un valor de 0.613 a 0.727 de área bajo la curva AUC, siendo propicios para la detección automática de diabéticos [6].
 - ✓ ES14: En este estudio: “Métodos de aprendizaje automático para detección de anomalías”, se pretende detectar las anomalías de un conjunto de datos usando los algoritmos de vecinos próximos, Random Forest y máquinas de soporte. Los métodos a tratar son el de covarianza robusta, el factor de outlier local, el de los bosques de aislamiento y el de las máquinas de vectores de soporte de una clase. Posteriormente se prueban los métodos en dos perspectivas: la primera en el uso de dos problemas sintéticos y la segunda en la aplicación de los métodos en conjuntos de datos reales, tanto para clasificación como para regresión. Adicionalmente, se ha aplicado un sistema de votación para la clasificación de los outliers, así se observa una mejoría variable en la predicción como en la clasificación. Se concluye que de los métodos empleados para los mejores resultados son los el del factor de outlier local, el de máquinas de vectores de soporte de una clase y el sistema de votación [76].

- ✓ ES15: En este estudio: “Valoración de riesgo cardiovascular mediante modelos de clasificación”, se trata de hacer una clasificación de pacientes que tienen algún evento cardiovascular como son personas de riesgo, sean estas hipertensos, diabéticos, con problemas de obesidad y tabaquismo. Además del modelo de Framingham, se utilizará los modelos de clasificación: árboles de decisión, regresión logística y random forest; de acuerdo a la tasa de clasificaciones correctas, eficiencia relativa y devianza [77].
- ✓ ES16: En este artículo: “Sistema e-health para capturar, transmitir y almacenar el estado de la presión arterial y otros datos relevantes, con el fin de calcular la probabilidad de padecer hipertensión arterial”, se implementa un sistema de análisis de datos, empleando Redes Bayesianas para determinar la probabilidad de sufrir hipertensión una persona. La metodología contiene: análisis de información, los requerimientos, la implementación y la validación del sistema. La utilización de la inteligencia artificial permitirá mejorar la calidad de vida de pacientes a través de la toma de decisiones por parte del médico [78].
- ✓ ES17: En este trabajo: “Diseño de un prototipo web interactivo basado en técnicas de machine learning y razonamiento probabilístico para pronosticar mediante los factores de riesgo la enfermedad de diabetes”, se realizó un modelo de clasificación de machine learning, este es regresión logística para determinar la enfermedad de diabetes de acuerdo a los factores de riesgo, se utilizó la metodología de investigación bibliográfica para el comportamiento de las variables usadas en el modelo probabilístico para convertirse en una herramienta inherente al médico y asegure un diagnóstico de prevención de diabetes [79].
- ✓ ES18: En este estudio: “Modelo basado en aprendizaje de máquina supervisado, para el análisis de datos cardiovasculares y su aplicación en el diagnóstico y pronóstico médico mediante el uso de una página web dinámica”, se utilizó la técnica de inteligencia artificial Redes Bayesianas para evaluar los factores de riesgo cardiovascular. Se utilizó dentro de la metodología investigación documental para las pruebas y verificación del comportamiento del modelo, aportando al diagnóstico de la patología médica [80].
- ✓ ES19: En este artículo: “Contribuciones de las técnicas machine learning a la cardiología. Predicción de reestenosis tras implante de stent coronario”, se desarrolló un modelo para predecir la aparición de reestenosis en pacientes con

infarto agudo de miocardio sometidos a Intervencionismo Coronario Percutáneo ICP con implante de un stent. La data corresponde a 263 pacientes, de los cuales 23 presentaron reestenosis, utilizando lenguaje Python, computación en la nube. Dentro de la metodología se utilizó el esquema de validación cruzada anidada, las curvas PR (Recuperación de Precisión), las curvas ROC (Área Bajo la Curva). El modelo extremely randomized trees dio un resultado de 0,96 para no reestenosis. Con un 18% más en relación al score de riesgo clásico TLR. Las variables más trascendentales fueron: diabetes, enfermedad coronaria, flujo TIMI post-ICP, plaquetas anormales, trombo post-ICP y colesterol anormal. Y se desarrolló una calculadora para estimar el riesgo individual de cada paciente y así el médico tome la mejor decisión para su tratamiento [81].

- ✓ ES20: En el artículo: “Extracción de información de Evoluciones Clínicas Digitales mediante técnicas de Machine Learning”, se emplea un modelo de machine learning para analizar los factores de riesgo cardiovascular de historias clínicas, empleando un procedimiento para el análisis y filtrado de las evoluciones principales en el modelo. Sus resultados demuestran la efectividad de los recursos empleados para la extracción de la información, obteniendo valores de ejecución idóneos [82].
- ✓ ES21: En este trabajo: “Desarrollo de una aplicación móvil para el monitoreo de pacientes con hipertensión arterial utilizando la tecnología ESP32”, se desarrolló un aplicativo web/móvil para brindar asistencia médica ambulatoria a pacientes hipertensos, utilizando análisis de regresión. Se realizó una revisión de literatura de journals, e-Books, tesis, páginas web, se utilizó el modelo espiral de Boehm para el monitoreo de pacientes con hipertensión, y se desarrolló el aplicativo con la emulación de datos en tiempo real de la presión arterial, utilizando la norma ISO 9241-11 para la evolución de criterios de usabilidad [83].
- ✓ ES22: En el artículo: “Minería de datos para la detección de factores de influencia en el test APGAR”, se aplicó técnicas de minería de datos para clasificar los valores de test de Apgar en los nacimientos en el Hospital Juan José de Urquiza en 2009 a 2017, se usaron los algoritmos: Class Balancer, SMOTE y Spread Sub Sampled y se aplicaron tres modelos de clasificación en árboles de decisión: J48, REP Tree y Random Tree. Los mejores resultados son de la combinación de Class Balancer y Random Tree. Los principales factores resultantes del modelo son la

aplicación de oxitócicos durante la labor de parto, la posición del niño al nacer, el peso y la edad gestacional del recién nacido [84].

- ✓ ES23: En este estudio: “Diabetes Mellitus Tipo 2 y sus efectos sobre procesos de Cognición Social”, se trató de analizar los posibles efectos de la diabetes sobre procesos de cognición social, comparando un conjunto de personas con DMT2 con el de un grupo de control en tareas de reconocimiento de emociones, empatía y toma de decisiones. Se evaluaron 40 personas del grupo experimental y 40 del grupo de control, se usó una prueba de escrutinio de estado de ánimo y otra de tamizaje cognoscitivo y para los de DMT2 una escala de actividades de autocuidado para la diabetes. Los resultados: el grupo de DMT2 tuvo un porcentaje menor de aciertos que el grupo de control en la prueba de reconocimiento de emociones; en el test de empatía, el grupo de DMT2 obtuvieron menor comprensión de situaciones de dolor accidentales y de situaciones neutrales. Y en el juego de cartas de Iowa no existieron diferencias significativas [85].
- ✓ ES24: En esta investigación: “Estudio de la relación del esfuerzo de corte con la presión en aneurismas cerebrales y la predicción del riesgo de ruptura usando herramientas de inteligencia artificial basado en datos morfológicos, fluidodinámicos y estructurales”, se tiene dos estudios: el primero se realizó simulaciones fluido dinámicas CFD para 60 aneurismas cerebrales humanas y así analizar el comportamiento del esfuerzo de corte de muro WSS promediado en el tiempo TAWSS con respecto al radio de aspecto AR de los aneurismas, bajo condiciones de hipotensión, hipertensión y presión normal. Su resultado fue un comportamiento lineal entre el TAWSS y el AR, y cuando aumentaban las diferencias de presión, el valor absoluto de la pendiente también incrementó. En el segundo estudio mostró que los parámetros morfológicos y fluido-mecánicos son los principales factores en la predicción del riesgo de ruptura de los aneurismas. Se usó 5 modelos de machine learning para predecir el riesgo de ruptura de los aneurismas, así el modelo de refuerzo adaptativo AdaBoost presenta el mayor valor del área bajo la curva AUC con un valor de 0,944 de características operativas del receptor ROC [86].
- ✓ ES25: En este artículo: “Desarrollo de una aplicación móvil (Appteroma) que permita estimar el riesgo de padecer Aterosclerosis utilizando algoritmos de

Machine Learning”, el trabajo de la aplicación radica en que el médico alimenta información básica del paciente durante una consulta y así la aplicación dará una predicción porcentual que permite ver el riesgo de sufrir aterosclerosis en años siguientes, este trabajo se encuadra en las Tecnologías de la Información y Comunicación en el área de la Salud y Machine Learning [87].

- ✓ ES26: En este artículo: “Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí”, se utilizó técnicas de aprendizaje automático para el análisis de imágenes cardíacas, la aportación de esta revolución tecnológica, limitaciones actuales, aspectos legales y sus aplicaciones iniciales en cardiología. Y este trabajo sirva para mejores diagnósticos y tratamientos [88].
- ✓ ES27: En este estudio: “Diseño y evaluación de un sistema de ayuda al diagnóstico para neuropatía diabética basado en la lectura de puntos de presión plantar y machine learning”, se desarrolló y evaluó el modelo usando una plantilla con sensores electrónicos y algoritmos de machine learning. El algoritmo diseñado tuvo una precisión de 0,94. Y tuvo gran aceptación de los endocrinólogos por ser un sistema fácil de usar y confiar, siendo un sistema de apoyo al diagnóstico [89].
- ✓ ES28: En este estudio: “Modelos de aprendizaje profundo para la predicción de comportamiento sedentario futuro”, se tasaron arquitecturas de aprendizaje profundo para predecir el comportamiento sedentario futuro de una persona a partir de los registros capturados de diferentes sensores disponibles en los dispositivos móviles actuales. Se analizaron usuarios con diferentes niveles de gasto energético obteniéndose resultados favorables de las arquitecturas planteadas [90].
- ✓ ES29: Este estudio: “Análisis neutrosófico para el diagnóstico de la hipertensión arterial a partir de un sistema experto basados en casos”, tiene como objetivo diagnosticar la hipertensión arterial HTA en una población de 455 hombres y 394 mujeres, de entre 18 y 78 años de edad. Los pacientes se clasificaron en normotensos, prehipertensos e hipertensos. Se usó técnicas multivariadas como el Análisis Discriminante y la Regresión Logística para jerarquizar los factores de riesgo de la HTA y obtener el grado de importancia de estos. Se utilizó la aplicación conjunta de un análisis de varianza ANOVA y el método TwoStep Cluster Analysis. Se aplicó un modelo de recomendación neutrosófica para las

semejanzas entre los casos analizados y los casos nuevos. Luego de obtener los resultados se comprueba el sistema experto validando su efectividad en el diagnóstico de la hipertensión arterial ante los factores de riesgo presentes en la población [91].

- ✓ ES30: En este trabajo: “Red neuronal para diagnóstico de hipertensión arterial”, se va a predecir la enfermedad utilizando diferentes tipos de clasificadores, exclusivamente redes neuronales, con una precisión de predicción de 99,8% [92].
- ✓ ES31: En el estudio: “Índice HOMA-IR como predictor de reducción de exceso de peso en pacientes con índice de masa corporal $IMC \geq 35$ kg/m² sometidos a gastrectomía vertical”, se evaluó la asociación del índice HOMA-IR con el porcentaje de sobrepeso perdido después de un año de una cirugía bariátrica, en pacientes mayores de 18 años entre el 2014 al 2017 en la Clínica Avendaño, Perú, usando regresión lineal y de Poisson. Los resultados son: 85,7% presentó resistencia a la insulina, y después de 1 año de la cirugía un 76,9% tuvo un porcentaje de sobrepeso perdido agradable [93].
- ✓ ES32: En este trabajo: “Métodos de aprendizaje supervisado para la predicción de diabetes: una revisión sistemática de literatura”, se identificó los mejores métodos para la predicción de la diabetes más significativos, destacándose: árbol de decisión y derivados. Se destacó la etapa de preprocesamiento en estos modelos para una mejor precisión [94].
- ✓ ES33: En el trabajo: “Determinación de cristatina c e identificación de la relación con la creatinina sérica en el diagnóstico de insuficiencia renal en pacientes con diabetes mellitus tipo 2”, se determinó la Cristatina C y la relación con la Creatina sérica para diagnosticar Insuficiencia Renal en pacientes con Diabetes Mellitus tipo 2. Se utilizó 30 registros de pacientes del Club de Diabéticos del Hospital Básico Pelileo, Ambato-Ecuador, donde el 73% de la población presentaron hiperglicemia, luego se estimó la TFG a partir de la Creatinina la cual estaba disminuida en el 100% y la TFG a partir de la Cristatina C estaba normal en el 50%. Utilizando el test estadístico Chi-cuadrado se verificó la correlación positiva de las variables, dando como resultado que la Cristatina C es mejor que la Creatinina sérica como marcador de enfermedad renal en pacientes diabéticos y así determinar el estadio del paciente [95].

- ✓ ES34: En este estudio: “Una guía conceptual para usar y entender Big Data en la investigación clínica”, se evidencia el uso del Big Data y la Ciencia de Datos para identificar los criterios para evaluar el impacto de estos métodos en la medicina y los desafíos que representa [69].
- ✓ ES35: En este trabajo: “Un enfoque de aprendizaje automático para la estimación de diagramas de fase para mezclas de lípidos de tres componentes”, se predijo a través de aprendizaje automático la alta temperatura de fusión de los lípidos y obtener diagramas de fase a partir de mezclas de lípidos. Se hizo uso de las redes neuronales, bosques aleatorios y máquinas de vectores de soporte. El modelo de bosques aleatorios generó un diagrama que coincidía estrechamente con los resultados de la literatura. Así se ahorró tiempo y recursos en el estudio de los lípidos [96].
- ✓ ES36: En este estudio: “Clasificación de factores sociodemográficos, de comportamiento de salud, de prevención y ambientales en la predicción de la salud cardiovascular en el vecindario: un enfoque de aprendizaje automático bayesiano”, se utilizó árboles de regresión aditiva bayesiana para clasificar los factores sociodemográficos, de comportamiento de salud, de prevención y ambientales para predecir la salud cardiovascular del vecindario. Se creó un conjunto de datos de los Centros para el Control y la Prevención de Enfermedades. Y los resultados verifican que los factores como personas con obesidad, que no tienen actividad física regular y el consumo de alcohol son los cinco predictores principales de riesgo cardiovascular del barrio [97].
- ✓ ES37: En este trabajo: “Prevención del riesgo cardiovascular/accidente cerebrovascular: un nuevo marco de aprendizaje automático que integra fenotipos basados en imágenes de ultrasonido carotídeo y sus armónicos con factores de riesgo convencionales”, se estratificó el riesgo de accidente cardiovascular usando bosques aleatorios con características como edad, sexo, hemoglobina Alc, azúcar en sangre en ayunas, lipoproteínas de baja densidad y colesterol de lipoproteínas de alta densidad, colesterol total y otras. El sistema usa 38 funciones. Se utilizó un enfoque de regresión logística para las características más significativas del clasificador Random Forest (bosques aleatorios). En los resultados se analizaron retrospectivamente las arterias carótidas de 202 pacientes. El clasificador RF mostró una mayor mejora en el área bajo la curva

AUC del 57%. Usando el clasificador de RF, las estadísticas de AUC para el sistema integrado AtheroRisk fueron más altas con $AUC=0,99$, valor $p<0,001$ respecto a AtheroRisk convencional con $AUC=0,63$, valor $p<0,001$. Así en AtheroRisk supera al sistema de aprendizaje automático convencional empleando Random Forest [98].

- ✓ ES38: En este artículo: “Un modelo predictivo preciso y dinámico para un sistema inteligente M-Health que utiliza aprendizaje automático”, se analizó los sistemas Emerging Mobile Health (M-Health), su modelo y arquitectura, tecnologías y aplicaciones, a través de enfoques estadísticos y de aprendizaje automático. Se propuso una arquitectura basada en Android para el registro de información de los pacientes, un modelo confiable en la nube, y así el modelo predictivo clasifique las enfermedades cardiovasculares. Los resultados son satisfactorios comparados con los modelos existentes por su precisión y especificidad [99].
- ✓ ES39: En este artículo: “Como el aprendizaje automático transformará la biomedicina”, se analizó como se puede transformar tres áreas de la biomedicina: diagnóstico clínico, tratamientos de precisión y monitoreo de la salud. Y esto constituye el futuro de la medicina [100].
- ✓ ES40: En esta investigación: “Imagen confocal asistida por aprendizaje automático de los sitios intracelulares de formación y almacenamiento de triglicéridos y ésteres de colesterilo”, se presentó una segmentación impulsada por la polaridad asistida por aprendizaje automático para cuantificar los sitios de biosíntesis de triglicéridos y ésteres de colesterilo en todos los orgánulos intracelulares, y así monitorear en tiempo real el recambio de estos lípidos no polares en células vivas, prediciendo y contrarrestando su progresión [101].
- ✓ ES41: En este trabajo: “Una revisión de las técnicas de aprendizaje automático en fotoplethismografía para la medición no invasiva sin manguito de la presión arterial”, se utilizó técnicas de procesamiento de señales, con aprendizaje automático para la monitorización continua y sin manguitos basados en fotoplethismograma (PPG) de la presión arterial, para hacer el seguimiento de los signos vitales de las personas con riesgo cardiovascular [102].
- ✓ ES42: En este artículo: “Investigación sobre el efecto del número de Womersley, las funciones de ECG y PPG para la estimación de la presión arterial sin manguito

mediante el aprendizaje”, se presentó un algoritmo basado en nuevas características en el dominio del tiempo para la monitorización continua de la presión arterial, importante en las unidades de cuidados intensivos. Las características fueron: el número de Womersley, QRS, intervalo QTc, SDI de las señales de ECG y PPG y técnicas de regresión empleadas para valorar la presión arterial de manera continua. Se emplea el algoritmo genético GA con el modelo de bosque aleatorio. Dentro de los resultados, el mejor conjunto de características óptimas de GA redujo el MAE de 13,20 a 9,54 mmHg, 9,91 a 5,48 mmHg, 7,71 a 3,37 mmHg para SBP, DBP y MBP respectivamente. Los resultados del algoritmo genético validan la importancia de la correlación de las características del ECG con la PA [103].

- ✓ ES43: En este trabajo: “Un nuevo método de estimación de la presión arterial basado en la clasificación de formas de onda oscilométricas mediante métodos de aprendizaje automático”, el conjunto de datos utilizados los atributos de la envolvente de onda oscilométrica se restaron para cada pulso, donde los atributos del latido del punto de presión sistólica es 1, y los atributos del latido de la presión diastólica es 2, los otros latidos se etiquetan con 0. Se volvió a etiquetar y luego de un total de 350 mediciones, 300 se dividieron entrenamiento y 50 en prueba. Se estimó tres algoritmos de clasificación, para cuatro métricas de rendimiento diferentes que son: error absoluto medio MAE, error cuadrático medio MSE, error cuadrático medio RSME y R los valores de la presión arterial sistólica PAS y la presión arterial diastólica PAD, se estimaron los algoritmos de k-vecino más cercano kNN, k-vecino más cercano ponderado WkNN y árboles en bolsas. Para la PAS con el algoritmo kNN, kNN ponderado y árboles en bolsas, los MAE obtenidos fueron 3.590, 3.520 y 4.499, respectivamente. Para los valores de DBP con el algoritmo kNN, kNN ponderado y árboles en bolsas, los MAE fueron 11.077, 11.032 y 13.069, respectivamente [104].
- ✓ ES44: En este trabajo: “Aprendizaje automático en el modelado de flujos cardiovasculares: predicción de la presión arterial a partir de datos de resonancia magnética de flujo 4D no invasivos utilizando redes neuronales informadas por la física”, se presentó un marco de aprendizaje automático para la síntesis perfecta de productos in-vivo no invasivo, técnicas de medición y modelos de dinámica de flujo computacional, derivados de los primeros principios físicos. Se mostró

como se utilizan los modelos unidimensionales de flujo pulsátil para restringir la salida de redes neuronales profundas de modo que sus predicciones satisfagan los principios de conservación de masa y momento. Entrenadas en datos clínicos ruidosos y dispersos de flujo y desplazamiento de la pared, estas redes devolvieron predicciones consistentes para la propagación de la onda de pulso de velocidad, presión y desplazamiento de la pared, sin emplear simuladores convencionales. Un postprocesamiento de estos resultados proporcionaría una forma eficaz de estimar los parámetros del modelo de Windkessel para la calibración de los modelos computacionales tradicionales [105].

- ✓ ES45: En este artículo: “Aprendizaje de propiedades físicas en escenas visuales complejas: una máquina inteligente para percibir la dinámica del flujo sanguíneo a partir de imágenes de angiografía por TC estática”, se propuso una red neuronal profunda para los valores de reserva fraccional de flujo FFR a partir de imágenes de angiografía coronaria por TC estática. Se usa una red neuronal recurrente estructurada en árbol RNN con un codificador de representación coronaria. El codificador captura información geométrica coronaria para el flujo sanguíneo. La RNN estructurada en árbol crea una dependencia espacial a larga distancia de la información del flujo sanguíneo dentro del árbol coronario. A 13000 árboles coronarios sintéticos y 180 árboles coronarios reales de pacientes clínicos, proyectaron valores del área bajo la curva ROC (AUC) de 0,92 y 0,93 bajo criterios clínicos [106].
- ✓ ES46: En este estudio: “Enfoque novedoso que combina aprendizaje máquina y modelado de dinámica de sistemas para la predicción del riesgo de eventos cardiovasculares secundarios y el efecto de las intervenciones”, de acuerdo a un marco biológico causal se usó un método nuevo que combina la dinámica del sistema de aprendizaje automático para modelar el riesgo cardiovascular secundario en pacientes ya infartados según datos reales. Los parámetros se estiman comparando la salida del modelo con los resultados binarios de eventos cardiovasculares adversos importantes MACE según los conjuntos de datos de entrenamiento, prueba y validación. Los resultados del área bajo la curva AUROC del modelo son de: 0,782 – 0,800. Esto facilitará la interacción entre el cardiólogo y el paciente [107].

- ✓ ES47: En este estudio: “Requisitos propuestos para la evaluación del aprendizaje automático relacionado con imágenes cardiovasculares (PRIME): una lista de verificación: revisada por el Consejo de innovación en el cuidado de la salud del American College of Cardiology”, su propósito fue tener una base científica en la investigación del aprendizaje automático para generar informes uniformes, trabajado por médicos y estadísticos bajo 7 conjuntos de requisitos. Así se obtiene una lista de elementos de informes como una lista de verificación de los pasos apropiados para garantizar la aplicación correcta de los modelos de aprendizaje automático y el informe coherente de las especificaciones y los resultados del modelo [108].
- ✓ ES48: En este estudio: “Integración de la puntuación STOP-BANG y los datos clínicos para predecir eventos cardiovasculares después de un infarto: un estudio de aprendizaje automático”, realizado a 124 pacientes con infarto de miocardio agudo, de los cuales la puntuación STOP-BANG, clasificó a 34 como bajo (27,4%), 30 como intermedio (24,2%) y 60 como alto (48,4%) pacientes con riesgo AOS que fueron seguidos en la hospitalización. El aprendizaje automático utilizó la selección e integración de características en 47 variables para identificar: muerte, arritmias ventriculares, fibrilación, angina recurrente, reinfarto, accidente cerebrovascular, empeoramiento de la insuficiencia cardíaca o shock cardiogénico. Dentro de los resultados hubo una proporción creciente de eventos cardiovasculares en los grupos de riesgo de AOS bajo, intermedio y alto ($p=0,005$). ML seleccionó 7 variables accesibles, su integración superó a todos los comparadores, con un área bajo la curva de 0,83, IC del 95%, 0,74 – 0,90; $p<0,01$ [109].
- ✓ ES49: En este estudio: “Uso del aprendizaje automático para predecir eventos cardiovasculares de un año en pacientes con miocardiopatía dilatada grave”, la finalidad es predecir los eventos cardiovasculares de 1 año en pacientes con miocardiopatía dilatada MCD grave usando aprendizaje automático ML y así ayudar a los galenos en la estratificación del riesgo y control del paciente. Se trabajó con el registro de 98 pacientes de dos centros. Se utilizaron en total 32 características de datos clínicos; se construyó un clasificador de Bayes y su rendimiento predictivo se evaluó utilizando el área bajo la curva AUC de las características operativas del receptor por validación cruzada de 10 veces. Los

resultados durante 1 año de seguimiento, 22 pacientes cumplieron el criterio del punto final del estudio, con una predicción de eventos cardiovasculares en pacientes con MCG grave con AUC, 0,887 y un intervalo de confianza del 95%, 0,813 – 0,961 [110].

- ✓ ES50: En este estudio: “Inteligencia artificial y aprendizaje automático en la atención de la salud cardiovascular”, se proporcionó una descripción general de la inteligencia artificial IA y el aprendizaje automático ML referente a la salud cardiovascular. Dentro de los resultados se incluyen interpretación automatizada de radiografías de tórax, electrocardiogramas, ecocardiogramas y angiografías, identificación de pacientes con insuficiencia cardíaca precoz mediante notas evaluadas por ML y predecir la mortalidad o complicaciones posteriores a procedimientos cardiovasculares quirúrgicos [111].
- ✓ ES51: En este estudio: “Detección de la gravedad de los trastornos cardiovasculares mediante el uso de características anatómicas del miocardio basado en un enfoque de máquina de aprendizaje extremo optimizado”, se integró las características anatómicas del miocardio del ventrículo izquierdo y la máquina de aprendizaje extremo optimizada ELM para la discriminación de personas con fracción de eyección anormal FE normal, leve, moderada y grave. Este trabajo se llevó a cabo en imágenes de resonancia magnética cardiovascular CMR de 104 personas en el segundo concurso anual de ciencia de datos de Kaggle. Respecto a los resultados el volumen del miocardio en el servicio de urgencias proporciona una mejor precisión de clasificación del 83,3% en comparación con otros. El marco BOELM dado proporcionó una precisión de clases múltiples más alta del 95,2% con todo el conjunto de características que ELM. Se logró una mejor distinción de sujetos sanos y anormales moderados que otros subgrupos [112].
- ✓ ES52: En este estudio: “Un enfoque basado en el aprendizaje automático para predecir el brote de enfermedades cardiovasculares en pacientes en diálisis”, la finalidad es predecir con puntualidad la muerte y las enfermedades cardiovasculares en pacientes dializados. Se utilizó dos conjuntos de datos, el primero un conjunto de datos italiano del Istituto di Fisiologia Clinica del Consiglio Nazionale delle Ricerche de Reggio Calabria y el segundo un conjunto de datos estadounidense del Instituto Nacional de Diabetes y Enfermedades Digestivas y Renales NIDDK. De cada uno se obtuvieron 5 conjuntos de datos.

Se probó diferentes algoritmos lineales como no lineales, finalmente se usó Support Vector Machine, utilizando el algoritmo de kernel SVC no lineal con RBF, optimizándolo con GridSearch. En los resultados se obtuvo una precisión del 95,2% para el conjunto de datos italiano y un 92,15% para el conjunto de datos americano, en un período de 2,5 años, respecto a la predicción de enfermedad isquémica del corazón [113].

- ✓ ES53: En este trabajo: “Técnicas de aprendizaje automático para predecir los resultados cardiovasculares en el Hospital en pacientes ancianos que presentan infarto de miocardio agudo”, se determinó las técnicas de aprendizaje automático ML para predecir enfermedades cardiovasculares intrahospitalarias, mortalidad y hemorragia mayor tras la intervención coronaria percutánea ICP en pacientes mayores de 70 años que presentan infarto de miocardio agudo. Se usó bosques aleatorios, regresión logística, máquina de vectores de soporte y redes neuronales artificiales. Todos los datos se dividieron aleatoriamente en 80% para entrenamiento y 20% para prueba. Para el modelo individual se utilizó validación cruzada de cinco veces; el rendimiento del modelo se evaluó con estadística C, métrica de calibración (puntuación de Brier), precisión y puntuación F1. En los resultados se identificaron 2038 pacientes con ICP con mediana de edad 77 años, 56% hombres. La tasa de mortalidad cardiovascular intrahospitalaria fue del 4,27% y de hemorragia mayor fue del 5,7%. Los tres predictores más significativos de mortalidad cardiovascular intrahospitalaria y hemorragia mayor fueron shock cardiogénico, infarto de miocardio y presencia de trombo [114].
- ✓ ES54: En esta investigación: “Poner en marcha el aprendizaje automático: aplicaciones en imágenes cardiovasculares”, su propósito fue ofrecer predicciones precisas en las imágenes cardíacas aplicando aprendizaje automático ML y así ayudar a los tratamientos específicos de pacientes y nuevas vías terapéuticas con una herramienta eficaz para el diagnóstico temprano y estratificación del riesgo para muerte prematura y discapacidad. El ML en imágenes cardiovasculares promete ser un instrumento importante y estudia una necesidad específica del paciente en la predicción precisa de eventos futuros y el descubrimiento de mecanismos moleculares manejables de la enfermedad [115].
- ✓ ES55: En este trabajo: “Usos y oportunidades del aprendizaje automático en la investigación de la hipertensión”, se utilizó la inteligencia artificial IA para

predecir resultados clínicos en big data, se incluyó cuatro ejemplos: en el primero, el aprendizaje profundo y la máquina de vectores de soporte SVM predijeron la aparición de eventos cardiovasculares con una precisión del 56% al 57%; en el segundo, en una base de datos de 378,256 pacientes, un algoritmo de red neuronal predijo la ocurrencia de eventos cardiovasculares durante 10 años de seguimiento con una sensibilidad del 68% y especificidad del 71%; en el tercero, un algoritmo de aprendizaje automático clasificó a 1.504.437 pacientes en presencia o ausencia de hipertensión con 51% de sensibilidad, 99% de especificidad y área bajo la curva del 87%; en el cuarto, se usaron biosensores y dispositivos portátiles para evaluar si una persona está en riesgo de desarrollar hipertensión mediante fotopletimografía para separar a las personas que tenían riesgo de desarrollar hipertensión con una sensibilidad superior al 80% y un valor predictivo positivo superior al 90%. Estos resultados se ajustaron según la demografía y los factores de riesgo tradicionales de la enfermedad aterosclerótica [116].

- ✓ ES56: El estudio: “Predicción dinámica de la mortalidad mediante técnicas de aprendizaje automático para casos cardiovasculares agudos”, presentó los resultados del uso de métodos de aprendizaje automático para la predicción temprana de la mortalidad de los pacientes cardiovasculares. Se analizó los datos de historias clínicas electrónicas de personas con síndrome coronario agudo, infarto y angina estable. Se desarrolló un enfoque para identificar los componentes del modelo y su conexión. Se analizaron datos dinámicos del sistema de información médica. Los resultados de pruebas de laboratorio simples como hemoglobina HGB, glóbulos rojos RBC, alanina transaminasa ALT, aspartato transaminasa AST, glucosa, plaquetas PLT, niveles de creatinina se utilizaron como predictores. Estos métodos de aprendizaje automático podrían ser usados considerablemente en la práctica clínica [117].
- ✓ ES57: En este estudio: “Aprovechamiento de las técnicas de aprendizaje automático para pronosticar el pronóstico del paciente después de una intervención coronaria percutánea”, se buscó determinar si el aprendizaje automático ML identifica mejor al paciente en riesgo de muerte y rehospitalización por insuficiencia cardíaca congestiva ICC después de una intervención coronaria percutánea ICP. El ML tiene la capacidad de identificar patrones no lineales complejos de conjuntos de datos, mejorando la predicción de

los modelos. Se utilizó el registro de 11.709 pacientes de PCI de Mayo Clinic, se examinaron 358 variables para identificar a las personas en riesgo de reingreso por ICC. En los resultados, el algoritmo predictivo identificó una cohorte de alto riesgo del 2% en pacientes que tenían una mortalidad hospitalaria del 45,5% en comparación con un riesgo del 2,1% para la población general (AUC= 0,925; intervalo de confianza del 95%: 0,92 a 0,93). Los predictores principales fueron: la edad avanzada, la insuficiencia cardíaca congestiva y el shock. Se identificó un grupo de alto riesgo que representaba el 1% de todos los pacientes con una rehospitalización por ICC de 30 días del 8,1% (intervalo de confianza del 95%: 6,3% a 10,2%). La regresión de bosque aleatoria superó a la regresión logística para predecir la readmisión por insuficiencia cardíaca congestiva de 30 días (AUC: 0,90 frente a 0,85; $p=0,003$; mejora neta en la reclasificación: 5,14%) y la muerte cardiovascular a los 180 días (AUC: 0,88 frente a 0,81; $p=0,02$; mejora neta de reclasificación: 0,02%). El ML fue eficaz para encontrar subgrupos con alto riesgo de mortalidad y readmisión posteriores al procedimiento [118].

- ✓ ES58: En este estudio: “Aprovechar los algoritmos de aprendizaje automático interpretables para predecir los resultados postoperatorios de los pacientes en dispositivos móviles”, se evidenció que el aprendizaje automático moderno ML ofrece modelos para capturar los factores de riesgo preoperatorios al riesgo general del paciente quirúrgico. Y así abordar deficiencias de los métodos tradicionales en un entorno clínico [119].
- ✓ ES59: En este estudio: “El aprendizaje automático mejora la predicción del riesgo de mortalidad después de la cirugía cardíaca: revisión sistemática y metaanálisis”, se realizó una revisión sistemática y un metaanálisis entre el modelo de aprendizaje automático ML y la regresión logística LR tradicional. La capacidad de discriminación se evaluó mediante el estadístico C. Las estadísticas C agrupadas y su intervalo de credibilidad del 95% para modelos ML y LR se obtuvieron usando Bayes. En los resultados, se identificaron 459 citas, donde 15 estudios cumplieron los criterios de inclusión para el análisis cuantitativo y cualitativo. Usando el mejor modelo de ML del estudio individual, las estimaciones meta analíticas mostraron que el ML se asoció con una estadística C significativamente más alta; mientras que cuando se seleccionaron algoritmos de ML individuales, se encontró una tendencia no significativa hacia una mejor

predicción para cada uno de los algoritmos de ML. Los algoritmos de ML obtuvieron mejor discriminación respecto a LR tradicional en la predicción de mortalidad después de la cirugía cardíaca [120].

- ✓ ES61: En este estudio: “Aprendizaje automático para predecir el riesgo a largo plazo de infarto de miocardio y muerte cardíaca en función del riesgo clínico, calcio coronario y tejido adiposo epicárdico: un estudio prospectivo”, se evaluó el rendimiento del aprendizaje automático ML, integrando parámetros clínicos con el calcio arterial coronario CAC y la cuantificación automatizada del tejido adiposo epicárdico EAT para la predicción del riesgo a largo plazo de infarto de miocardio IM y muerte cardíaca en personas asintomáticas. Este estudio incluyó a 1.912 personas asintomáticas. El volumen y la densidad de EAT se cuantificó por un método de ML totalmente automatizado. El refuerzo de gradiente extremo se ML se entrenó usando covariables clínicas, mediciones de panel de lípidos en plasma, factores de riesgo, CAC, calcio aórtico y medidas de EAT automatizadas. En el seguimiento medio se produjeron 76 eventos de infarto de miocardio y / o muerte cardíaca. ML obtuvo un área bajo la curva AUC mayor que el riesgo de enfermedad cardiovascular aterosclerótica ASCVD y la puntuación CAC para predecir eventos (ML: 0,82; ASCVD: 0,77; CAC: 0,77, $p < 0,05$ para todos). Personas con una puntuación ML más alta tenían un alto riesgo de sufrir eventos (HR; 10,38, $p < 0,001$), las relaciones persistieron en el análisis multivariable, incluidas las medidas de riesgo ASCVD y CAC (HR: 2,94, $p = 0,005$). La edad, el riesgo de ASCVD y el CAC fueron importantes para el pronóstico de ambos sexos [121].
- ✓ ES62: En este estudio: “El aprendizaje automático revela que los esfingolípidos séricos son biomarcadores independientes del colesterol de la enfermedad de las arterias coronarias”, se realizó una lipidómica dirigida en muestras de suero en pacientes con enfermedad arterial coronaria familiar EAC (n=462) y controles basados en la población (n=212) para explorar la relación entre esfingolípidos séricos y CAD usando aprendizaje automático ML imparcial para identificar especies de esfingolípidos asociados con CAD. En los resultados, casi todos los esfingolípidos medidos (n=30 de 32) estaban significativamente elevados en sujetos con CAD en comparación con mediciones de controles de población. Se generó una nueva puntuación de riesgo EAC que incluye esfingolípidos que

delimita a los pacientes con EAC de forma independiente y más eficaz que los biomarcadores de ECV clínicos convencionales, incluidos el colesterol LDL y los triglicéridos séricos. Se concluye que este estudio validó las ceramidas séricas como biomarcadores candidatos de ECV y así los paneles de esfingolípidos deben considerarse medidas de CVD [122].

- ✓ ES63: En este estudio: “La lipidómica de espectrometría de masas no dirigida identifica la correlación entre las esfingomielinas séricas y el colesterol plasmático”, se investigó las relaciones de los cambios de los esfingolípidos séricos, la regulación del metabolismo de las lipoproteínas de baja densidad LDL y lipoproteínas de alta densidad HDL y los endofenotipos de la enfermedad de las arterias coronarias CAD. Se realizó cromatografía líquida no dirigida - espectrometría de masas (UPLC-MS) lipidómica de muestras de suero de pacientes pertenecientes a un estudio transversal sobre la base de la ausencia o presencia de enfermedad arterial coronaria EAC definida angiográficamente y caracterizados extensamente por clínica y fenotipos bioquímicos. Dentro de los resultados, 2998 características espectrales detectadas en las muestras de suero, 1328 características metabólicas se correlacionaron significativamente con al menos uno de los fenotipos clínicos o bioquímicos medidos en la cohorte. Se evidenció asociaciones significativas entre 34 señales de metabolitos de un conjunto de esfingomielinas y el colesterol HDL sérico. Muchas de estas asociaciones de metabolitos también se observaron con las concentraciones séricas de LDL y colesterol total, pero no tanto con los triglicéridos séricos. Se concluyó que de los pacientes con EAC, los esfingolípidos se correlacionan directamente con los niveles séricos de lipoproteínas y colesterol total [123].
- ✓ ES64: En este estudio: “Valor predictivo de la proteómica dirigida para la morfología de la placa coronaria en pacientes con sospecha de enfermedad arterial coronaria”, se investigó la capacidad de la proteómica dirigida para predecir la placa de alto riesgo o ausencia de aterosclerosis coronaria en pacientes con enfermedad arterial coronaria EAC definida por angiografía por tomografía computarizada coronaria CCTA. Se usaron niveles plasmáticos de 358 proteínas para generar modelos de aprendizaje automático ML. Su rendimiento se probó en comparación a un modelo clínico de características clínicas generalmente disponibles y biomarcadores convencionales. Se utilizó un total de 196 pacientes

con niveles de proteína analizables, identificando un subconjunto de 35 proteínas que predice la presencia de placas de alto riesgo. El modelo de ML tuvo un rendimiento aceptable con un área bajo la curva $AUC=0.79 \pm 0.01$, superando a la predicción convencional así: $AUC=0.65 \pm 0.04$, $p<0.05$. Así mismo, un subconjunto diferente de 34 proteínas predijo la ausencia de CAD con $AUC=0.85 \pm 0.05$, nuevamente superando a la predicción tradicional así: $AUC=0.70 \pm 0.04$, $p<0.05$. Este estudio se realizó con el apoyo de HeartFlow Inc. y la Red del Área de Investigación Europea de Enfermedades Cardiovasculares [124].

- ✓ ES65: En este estudio: “Factores de riesgo de enfermedad arterial coronaria en pacientes jóvenes con angina de pecho estable”, se realizó una revisión de perfiles de lípidos cardíacos en modelos de infarto de miocardio, obesidad y diabéticos, y miocardiopatía dilatada de enfermedad cardiovascular ECV mediante métodos de análisis de espectrometría de masas [125].
- ✓ ES66: En este estudio: “Predecir las complicaciones de la diabetes mellitus mediante algoritmos avanzados de aprendizaje automático”, se pretendió predecir si los pacientes con diabetes mellitus tipo 2 DM2 de un Hospital de California, desarrollarían 10 complicaciones seleccionadas para ralentizar su desarrollo. Se utilizó aprendizaje automático ML en comparación con los modelos tradicionales; se usó bosques aleatorios y redes neuronales RNN. En los resultados, los mejores se obtuvieron con el modelo RNN GRU, seguido del modelo RNN LSTM. El modelo RNN GRU estuvo entre el 73% para infarto de miocardio, 83% en cardiopatía isquémica crónica, en comparación con los modelos tradicionales estuvo entre el 66% y el 76% respectivamente [126].
- ✓ ES67: En este trabajo: “Predecir el riesgo a 10 años de complicaciones en los órganos terminales de la diabetes tipo 2 con y sin cirugía metabólica: un enfoque de aprendizaje automático”, el propósito fue construir y validar modelos de predicción para estimar el riesgo de complicaciones de órganos en pacientes con diabetes tipo 2 y obesidad para aportar a la toma de decisiones de los médicos especialistas, respecto a la cirugía metabólica. Se obtuvo el registro de 2287 pacientes de la Clínica de Cleveland con índice de masa corporal IMC mayor a 30 kg/m² y diabetes tipo 2. Se crearon y validaron modelos de regresión multivariable y de bosques aleatorios RF para predecir el riesgo a 10 años. Los modelos de predicción se programaron para teléfonos inteligentes y basadas en la

web de complicaciones de las diabetes individualizadas IDC para uso clínico. Los resultados demostraron la capacidad de discriminación basada en el área bajo la curva (1=discriminación perfecta y 0,5=probabilidad) a los 10 años en los grupos quirúrgico y no quirúrgico, mortalidad por todas las causas (0,79 y 0,81), eventos coronarios (0,66 y 0,67), insuficiencia cardíaca (0,73 y 0,75) y nefropatía (0,73 y 0,76). Se concluyó que los puntajes de riesgo de IDC podrán proporcionar información de riesgo personalizada basada en evidencia para pacientes con diabetes tipo 2 y obesidad sobre los resultados cardiovasculares futuros y la mortalidad con y sin cirugía metabólica según su estado actual de obesidad, diabetes y afecciones cardiometabólicas relacionadas [127].

- ✓ ES68: En este estudio: “Explorando factores de riesgo de insuficiencia cardíaca a través del aprendizaje automático”, se ha examinado las correlaciones entre infartos de miocardio y los índices de la bolsa. Se analizó los datos diarios de infartos agudos de miocardio producidos en Cataluña, España del año 2010 al 2016. Mediante las técnicas de aprendizaje automático se implementó un sistema de predicción de infartos, se realizó una comparativa de estas técnicas para obtener el mejor desempeño de la data [128].
- ✓ ES69: En este trabajo: “Los niveles séricos de ácido úrico y colesterol unido a lipoproteínas de baja densidad son predictores independientes de enfermedad arterial coronaria en pacientes indios asiáticos con diabetes mellitus tipo 2”, se identificó los predictores de enfermedad arterial coronaria EAC en 50 pacientes asiáticos, mayores de 40 años, con diabetes mellitus tipo 2 DMT2. Se evaluó los niveles de glucosa plasmática posprandial y en ayunas, el perfil lipídico, urea en sangre, la creatinina sérica y los niveles séricos de ácido úrico. Dentro de los resultados, los predictores más significativos fueron: el índice de masa corporal IMC, la relación cintura-cadera, ácido úrico sérico, colesterol total sérico, colesterol de lipoproteínas de baja densidad LDL y triglicéridos. Se concluyó que el colesterol LDL en suero y la hiperuricemia sirven como predictores independientes de arteriopatía coronaria CAD entre los pacientes seleccionados con DMT2 [129].
- ✓ ES70: En este estudio: “Partículas similares a remanentes y enfermedad de las arterias coronarias en la hipercolesterolemia familiar”, se investigó la asociación entre el colesterol de partículas remanentes RLP-C y la presencia de enfermedad

arterial coronaria EAC en pacientes con hipercolesterolemia familiar HF. Se trabajó con un registro de 282 pacientes con HF, donde se midió los niveles de RLP-C, las características basales, niveles de lípidos, otros factores de riesgo convencionales, la presencia de EAC y los niveles séricos de RLP-C. En los resultados se encontró que los niveles séricos de RLP-C se correlacionaron significativamente con los niveles séricos de triglicéridos TG (r de Pearson = 0,631, $p < 0,001$). Se observó que una mayor proporción de personas en los terciles superiores de RLP-C sérico tenían un mayor número de arterias coronarias enfermas ($p < 0,001$) para la tendencia de enfermedad multivaso. El análisis de regresión logística, después de ajustar por edad, sexo, hipertensión, diabetes, tabaquismo, colesterol de lipoproteína de baja densidad LDL-C reveló que la RLP-C se asoció significativamente con arteriopatía coronaria CAD, sin cambio la adición de los niveles séricos de TG en el modelo de regresión logística anuló esta asociación, en el caso de LDL-C se asoció independientemente con la EAC. Se concluyó que los niveles séricos de RLP-C se asociaron significativamente con la amenaza de EAC en dichos pacientes con HF [130]

1	HISTORIN/SEXO	EDAD	DMT7	HITA	COL-T	TRIGLII	HDL-C	LDL-C	COL_NO	RAZA	PAS	PAD	IMC	FUMAI	ALCOHC	TTTO_EST	TTTO_AL	TTTO_A3	SMY	IC	HOSPITA	MUERTE_CV	Col-t	_Normi	Trigliceridos	Hdl-C	_Normi	Ldl-C	_N
2	41781 AF	61	1	1	275	235	46	183	229	1	121	62	30	NO	NO	SI	NO	NO	0	0	0	0	0,4493	0,2423	0,6571	0,			
3	201302 AF	41	1	1	259	139	48	192	211	1	116	69	29	NO	NO	NO	SI	NO	0	0	0	0	0,4232	0,1433	0,6857	0,			
4	7490 AM	44	1	1	186	194	49	99	137	1	114	62	37	NO	NO	SI	SI	NO	0	0	0	0	0,3039	0,2000	0,7000	0,			
5	6436 AF	37	1	0	153	129	47	107	106	1	144	83	39	NO	NO	NO	SI	NO	0	0	0	0	0,2500	0,1330	0,6714	0,			
6	48925 AF	63	1	1	360	317	65	128	295	1	127	90	35	NO	NO	NO	SI	NO	0	0	1	0	0,5882	0,3268	0,9286	0,			
7	54023 AM	82	0	1	147	119	44	68	103	1	100	80	17	NO	NO	SI	SI	NO	0	1	1	0	0,2402	0,1227	0,6286	0,			
8	238347 AF	60	0	0	152	214	44	101	108	1	115	65	32	NO	NO	SI	NO	SI	0	0	0	0	0,2484	0,2206	0,6286	0,			
9	46667 AF	49	0	1	195	114	55	77	140	1	154	89	40	NO	NO	SI	NO	NO	0	0	0	0	0,3186	0,1175	0,7857	0,			
10	18804 AF	72	0	1	271	132	47	198	224	1	135	80	34	NO	NO	NO	SI	NO	0	0	0	0	0,4428	0,1361	0,6714	0,			
11	13272 AM	39	1	0	206	139	47	135	159	1	124	83	30	NO	NO	SI	NO	NO	0	0	0	0	0,3366	0,1433	0,6714	0,			
12	130627 AF	81	0	1	158	368	40	129	118	1	93	50	21	NO	NO	SI	SI	NO	1	0	1	1	0,2582	0,3794	0,5714	0,			
13	16530 AM	38	1	0	246	153	48	149	198	1	120	70	30	NO	NO	SI	NO	NO	0	0	1	0	0,4020	0,1577	0,6857	0,			
14	25678 AM	48	1	1	302	280	54	110	248	1	130	70	28	NO	NO	NO	SI	NO	0	0	0	0	0,4935	0,2887	0,7714	0,			
15	20765 AM	50	1	0	214	284	40	129	174	1	100	80	33	NO	NO	SI	NO	NO	0	0	0	0	0,3497	0,2928	0,5714	0,			
16	56626 AF	63	1	1	179	205	40	106	139	1	120	80	34	NO	NO	NO	SI	NO	0	0	0	0	0,2925	0,2113	0,5714	0,			
17	15519 AF	31	1	0	234	189	44	168	190	1	123	71	29	NO	NO	SI	NO	NO	0	0	0	0	0,3824	0,1948	0,6286	0,			
18	32374 AF	60	1	1	167	99	55	102	112	1	120	60	28	NO	NO	SI	NO	NO	0	0	1	0	0,2729	0,1021	0,7857	0,			
19	94764 AF	65	1	0	200	205	50	142	150	1	109	55	33	NO	NO	SI	NO	NO	0	0	0	0	0,3268	0,2113	0,7143	0,			
20	90089 AF	84	0	1	193	134	54	112	139	1	110	62	25	NO	NO	SI	SI	NO	0	0	0	0	0,3154	0,1381	0,7714	0,			
21	12743 AM	52	1	1	141	184	43	92	98	1	120	80	30	NO	NO	NO	SI	SI	0	0	0	0	0,2304	0,1897	0,6143	0,			
22	5666 AM	57	1	1	187	482	45	112	142	1	155	86	32	NO	NO	SI	SI	NO	0	0	0	0	0,3056	0,4969	0,6429	0,			
23	15531 AM	62	0	1	132	108	48	73	84	1	124	68	30	NO	NO	SI	SI	NO	0	1	0	0	0,2157	0,1113	0,6857	0,			
24	160241 AM	78	0	1	145	459	44	190	101	1	107	77	22	NO	NO	NO	SI	SI	0	1	1	0	0,2369	0,4732	0,6286	0,			
25	39099 AM	61	1	0	156	280	43	90	113	1	110	73	31	NO	NO	NO	SI	NO	0	0	1	0	0,2549	0,2887	0,6143	0,			
26	18109 AF	60	1	0	175	136	60	106	115	1	100	67	27	NO	NO	NO	NO	NO	0	0	0	0	0,2859	0,1402	0,8571	0,			
27	21922 AF	63	1	1	188	331	46	87	142	1	109	72	29	NO	NO	SI	SI	NO	0	0	0	0	0,3072	0,3412	0,6571	0,			
28	17667 AF	40	1	0	208	268	52	154	156	1	104	84	36	NO	NO	NO	NO	NO	0	0	0	0	0,3399	0,2763	0,7429	0,			
29	120865 AF	23	1	0	160	128	55	80	105	1	117	65	35	NO	NO	NO	SI	NO	0	0	0	0	0,2614	0,1320	0,7857	0,			
30	21389 AF	59	1	0	180	90	65	97	115	1	121	60	31	NO	NO	NO	NO	SI	NO	0	0	1	0	0,2941	0,0928	0,9286	0,		
31	10968 AM	48	1	0	233	195	42	172	191	1	118	85	22	NO	NO	SI	NO	NO	0	0	0	0	0,3807	0,2010	0,6000	0,			
32	71737 AM	46	1	0	167	178	45	122	122	1	120	60	23	NO	NO	SI	SI	NO	0	0	0	0	0,2729	0,1835	0,6429	0,			
33	1871 AF	43	1	0	212	111	62	128	150	1	104	62	22	NO	NO	SI	NO	NO	0	0	0	0	0,3464	0,1144	0,8857	0,			

Figura 39: Base de datos general final

Anexo 4: Documento de Especificación de Requerimientos.

Especificación de Requerimientos de Software

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos

2021

Ficha del documento

Fecha	Revisión	Autor	Verificado
17/09/2021		Roger Torres	

1. Introducción

El presente documento es una Especificación de Requerimientos de Software (ERS) para la solución informática CardioZT vía aplicación web. Esta especificación se ha estructurado basándose en las directrices dadas por el estándar IEEE Práctica Recomendada para Especificaciones de Requisitos Software ANSI/IEEE 830, 1998.

1.1. Propósito

El presente documento tiene como propósito definir las especificaciones funcionales, no funcionales para el desarrollo de la Solución informática CardioZT vía aplicación web que a la vez permitirá dar un score del riesgo cardiovascular (DMT2 e HTA) a través del perfil lipídico.

1.2. Alcance

Esta especificación de requisitos está dirigida a los usuarios del sistema, en la cual la aplicación estará enfocada principalmente en diseñar y desarrollar un sistema que posibilite al médico, administrar eficientemente el control del perfil lipídico y su relación con el riesgo cardiovascular de sus pacientes con DMT2 e HTA, para una mejor toma de decisiones.

1.3. Personal Involucrado

Tabla 20: Responsable del proyecto a desarrollar

Nombre	Francisco Javier Álvarez Pineda
Rol	Director del Trabajo de Titulación
Categoría Profesional	Docente de la carrera Ingeniería en Sistemas
Responsabilidad	Supervisar y asesorar el desarrollo del Trabajo de Titulación.
Información de contacto	fjalvarez@unl.edu.ec

Tabla 21: Director del Trabajo de Titulación

Nombre	Roger Alexander Torres Yaguana
Rol	Analista, diseñador y programador
Categoría Profesional	Estudiante de la carrera Ingeniería en Sistemas
Responsabilidad	Análisis de información, diseño y programación del CardioZT
Información de contacto	roger.torres@unl.edu.ec

1.4. Definiciones, acrónimos y abreviaturas

Tabla 22: Definiciones, Acrónimos y Abreviaturas

Nombre	Descripción
Usuario	Persona que utilizará el sistema para la predicción del riesgo cardiovascular y el cálculo del perfil lipídico
CardioZT	Nombre del sistema a desarrollar
ERS	Especificación de Requerimientos de Software
RF	Requerimiento Funcional
RNF	Requerimiento No Funcional
Django	Framework de desarrollo para aplicación web
virtualenv	Entorno virtual
Pandas	Biblioteca para manipulación y análisis de datos
Numpy	Librería para cálculo numérico de big data
Scikit-learn	Biblioteca de aprendizaje automático

1.5. Referencias

Título del Documento	Referencia
Estándar IEEE 830 -1998	IEEE

1.6. Resumen

Este documento consta de tres secciones. En la primera sección se realiza una introducción al mismo y se proporciona una visión general de la especificación de recursos del sistema.

En la segunda sección del documento se realiza una descripción general del sistema, con el fin de conocer las principales funciones que éste debe realizar, los datos asociados y los factores, restricciones, supuestos y dependencias que afectan al desarrollo, sin entrar en excesivos detalles.

Por último, la tercera sección del documento es aquella en la que se definen detalladamente los requisitos que debe satisfacer el sistema.

2. Descripción General

2..1 Perspectiva del producto

La Solución informática CardioZT será un producto diseñado para trabajar en un entorno web, lo que permitirá su utilización de forma rápida y eficaz. Este sistema será un producto diseñado para trabajar desde cualquier lugar, además se integrará con los diferentes sistemas que ya se encuentran funcionando dentro del hospital, de esta manera se espera lograr una mejor atención al paciente.

2.2 Funcionalidad del Producto

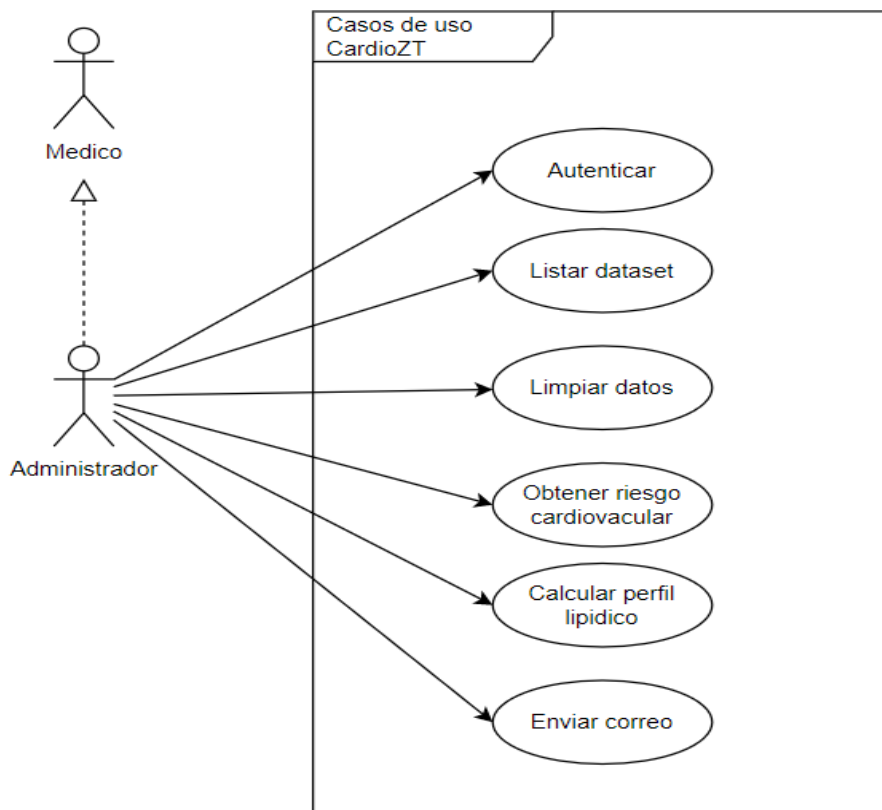


Figura 40: Diagrama de Casos de Uso

2.3. Características de los usuarios

Tabla 23: Características del Usuario Administrador

Tipo de Usuario	Administrador
Formación	Persona natural con conocimientos básicos en informática
Actividades	Administración y manejo del sistema en general

Tabla 24: Características del Usuario Contribuyente

Tipo de Usuario	Médico
Formación	Persona natural con conocimientos en Medicina
Actividades	Consulta de la predicción del riesgo cardiovascular y el cálculo del perfil lipídico.

2.4 Restricciones

- ✓ Lenguajes y tecnologías a usar: Python, HTML, CSS, JavaScript.
- ✓ Interfaz para ser utilizada en un entorno web.
- ✓ El sistema desarrollado consumirá los datos desde un archivo estático csv.
- ✓ El servidor deberá estar construido mediante la arquitectura de software Modelo Vista Template (MVT).
- ✓ El sistema deberá responder a las peticiones realizadas por el usuario.

2.5 Suposiciones y dependencias

- ✓ Se asume que los requerimientos aquí detallados son los necesarios para el correcto funcionamiento del sistema.

3. Requisitos específicos

3.1 Requisitos comunes de las interfaces

3.1.1. Interfaces de usuario

Todos los módulos del sistema contarán con una interfaz amigable, deberá estar conformada por botones, campos de texto, listas y notificaciones que permitirán al usuario navegar con facilidad y a su vez ser visualizada desde cualquier navegador de internet.

i. Interfaces de hardware

Para un correcto funcionamiento del sistema, es necesario contar con una computadora, Tablet o celular en óptimas condiciones con las siguientes características:

- ✓ Procesador de 1.30Ghz o superior.
- ✓ Adaptador de red.
- ✓ Memoria RAM de 2GB o superior.
- ✓ Ratón
- ✓ Teclado

ii. Interfaces de Software

- ✓ Sistema Operativo: Linux, Windows o MacOS
- ✓ Explorador: Mozilla Firefox, Google Chrome, Microsoft Edge, Opera.

iii. Interfaces de comunicación

La comunicación que se dará mediante los protocolos de internet más comunes, como por ejemplo el protocolo TCP/IP y HTTP.

b. Requisitos Funcionales

3.2.1. Requisito Funcional 1

Tabla 25: Requerimiento Funcional 1

Código	RF01
Nombre	Autenticar Usuario
Propósito	Iniciar sesión en la aplicación web
Descripción	Como médico administrador del sistema necesito autenticarme con el usuario y contraseña.
Entrada	Credenciales
Salida	Mostrar la página principal de la aplicación web
Prioridad	Media

3.2.2. Requisito Funcional 2

Tabla 26: Requerimiento Funcional 2

Código	RF02
Nombre	Cargar los datos cuantitativos de laboratorio del paciente
Propósito	Leer los datos del conjunto total de registros
Descripción	Como médico administrador del sistema necesito cargar los datos cuantitativos de laboratorio del paciente
Entrada	Nombre de dataset, descripción y fecha
Salida	Lista de datasets del sistema
Prioridad	Alta

3.2.3. Requisito Funcional 3

Tabla 27: Requerimiento Funcional 3

Código	RF03
Nombre	Limpiar los datos cuantitativos del conjunto total de pacientes
Propósito	Disponer de los datos limpios del dataset
Descripción	Como médico administrador del sistema necesito limpiar el total de datos de los archivos csv
Entrada	Nombre del dataset
Salida	Obtención de datos limpios del dataset
Prioridad	Alta.

3.2.4. Requisito Funcional 4

Tabla 28: Requerimiento Funcional 4

Código	RF04
Nombre	Mostrar los datos normalizados en base a métricas del sistema
Propósito	Obtener los datos para la administración del modelo predictivo
Descripción	Como médico administrador del sistema necesito seleccionar las variables del dataset a utilizar para hacer el pronóstico del riesgo cardiovascular
Entrada	Variables independientes y dependiente del modelo. Porcentaje de entrenamiento
Salida	Modelo creado con regresión lineal múltiple
Prioridad	Alta

3.2.5. Requisito Funcional 5

Tabla 29: Requerimiento Funcional 5

Código	RF05
Nombre	Ingresar datos para predecir el riesgo cardiovascular
Propósito	Obtener los valores de las variables para la predicción
Descripción	Como médico administrador del sistema necesito ingresar los valores para el riesgo cardiovascular.
Entrada	Valores de los campos de las variables
Salida	Campos validados en variables categóricas
Prioridad	Alta

3.2.6. Requisito Funcional 6

Tabla 30: Requerimiento Funcional 6

Código	RF06
Nombre	Mostrar el porcentaje de riesgo cardiovascular del paciente
Propósito	Insertar las lecturas por cada usuario registrado del sistema.
Descripción	Como médico administrador del sistema necesito visualizar el resultado del riesgo cardiovascular del paciente.
Entrada	Variables de Sexo, Edad, Dmt2, Col-T, Hdl-C, Pas, Fumador, Alcohol.
Salida	Mostrar en pantalla el porcentaje del riesgo cardiovascular
Prioridad	Alta.

3.2.7. Requisito Funcional 7

Tabla 31: Requerimiento Funcional 7

Código	RF07
Nombre	Mostrar gráficos estadísticos de los datos cargados en el sistema
Propósito	Visualizar la forma de los datos cargados en el dataset
Descripción	Como médico administrador del sistema necesito ver las estadísticas de los datos cargados en el sistema
Entrada	Valores del conjunto total de datos.
Salida	Gráficas del conjunto total de pacientes en género, en enfermedad de diabetes e hipertensión, rangos de edad, tratamientos antihipertensivos
Prioridad	Media.

3.2.8. Requisito Funcional 8

Tabla 32: Requerimiento Funcional 8

Código	RF08
Nombre	Ingresar datos para calcular el perfil lipídico
Propósito	Visualizar las variables que corresponden al perfil lipídico del paciente seleccionado.
Descripción	Como médico administrador del sistema necesito ingresar las variables para el perfil lipídico del paciente.
Entrada	Variables de Col-T, Sexo, Hdl-C, No-Hdl-C, Ldl-C y Triglicéridos
Salida	Campos llenados correctamente tanto para hombre como para mujer.
Prioridad	Media

3.2.9. Requisito Funcional 9

Tabla 33: Requerimiento Funcional 9

Código	RF09
Nombre	Mostrar el nivel del perfil lipídico del paciente.
Propósito	Verificar el estado del colesterol del paciente
Descripción	Como médico administrador del sistema requiero ver el nivel del perfil lipídico que maneja el paciente en prevención secundaria
Entrada	Variables para el perfil lipídico
Salida	Niveles del perfil lipídico en dos categorías para cada uno de sus componentes: en meta y fuera de meta de control.
Prioridad	Alta

3.2.1.0. Requisito Funcional 10

Tabla 34: Requerimiento Funcional 10

Código	RF10
Nombre	Enviar mediante correo los resultados obtenidos del perfil lipídico
Propósito	Guardar la valoración del perfil del paciente
Descripción	Como médico administrador del sistema necesito enviar por correo electrónico el resultado del nivel del perfil lipídico del paciente.
Entrada	Resultados del cálculo del perfil lipídico
Salida	Mensaje de envío satisfactorio de los resultados del perfil lipídico del paciente analizado.
Prioridad	Alta

c. Requisitos No funcionales

Tabla 35: Requerimiento No Funcional 1

Código	RNF01
Nombre	Look and Feel
Descripción	El sistema web debe tener un interfaz amigable para el usuario y brindar información consistente.
Prioridad	Alta

Tabla 36: Requerimiento No Funcional 2

Código	RNF02
Nombre	Robustez
Descripción	El sistema deberá sea capaz de manejar toda la información de una manera rápida y precisa.
Prioridad	Media

Tabla 37: Requerimiento No Funcional 3

Código	RNF03
Nombre	Confidencialidad
Descripción	La información de los usuarios será manipulada de forma segura y con fines administrativos.
Prioridad	Alta

Tabla 38: Requerimiento No Funcional 4

Código	RNF04
Nombre	Seguridad
Descripción	El sistema manejará las contraseñas de forma encriptada.
Prioridad	Alta

Tabla 39: Requerimiento No Funcional 5

Código	RNF05
Nombre	Disponibilidad
Descripción	El sistema debe garantizar su disponibilidad durante el tiempo que se realiza el cobro del servicio.
Prioridad	Alta.

Tabla 40: Requerimiento No Funcional 6

Código	RNF06
Nombre	Desarrollado mediante una metodología
Descripción	El sistema será desarrollado con la metodología de desarrollo ágil Extreme Programming (XP).
Prioridad	Media

Anexo 5: Definición Siglas y Abreviaturas.

Tabla 41: Definiciones, Siglas y Abreviaturas

Nombre	Descripción
XP	Programación Extrema (Extreme Program)
Django	Framework de desarrollo web Frontend basado en Python
Sklearn	Conjunto de rutinas para análisis predictivo
Pandas	Biblioteca de python para el manejo y análisis de datos
Numpy	Librería para grandes volúmenes de datos
Linear Regression	Modelo matemático de Regresión lineal
virtualenv	Entorno de trabajo de la aplicación en Python
ORM	Mapeo Relacional de Objetos (Object Relational Mapper)
ERS	Especificación de Requerimientos del Software
TT	Trabajo de Titulación
AA	Aprendizaje Automático
IA	Inteligencia Artificial
ECV	Enfermedad Cardiovascular Aterosclerótica
ASCVD	Enfermedad Aterosclerótica
SCORE	Systematic Coronary Risk Estimation
ERC	Enfermedad Renal Crónica
EAS	Sociedad de Aterosclerosis Europea
CAC	Calcio Arterial Coronario
HF	Hipercolesterolemia Familiar
SCA	Síndrome Coronario Agudo
HTA	Hipertensión Arterial

PAS	Presión Arterial Sistólica
PAD	Presión Arterial Diastólica
IMC	Índice de Masa Corporal
DMT2	Diabetes Mellitus Tipo 2
COL-T	Colesterol Total
TRIG	Triglicéridos
HDL-C	Lipoproteína de Alta Densidad
LDL-C	Lipoproteína de Baja Densidad
COL NO-HDL	Colesterol No HDL
TTO. ESTATINAS	Tratamiento con Estatinas
TTO. ANTI-HTA	Tratamiento Antihipertensivo
TTO. ASPIRINA	Tratamiento con Aspirina
IC	Insuficiencia Cardíaca

Anexo 6: Arquitectura del Sistema.

Arquitectura del Sistema

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos

1. Introducción

En el presente documento se detalla la arquitectura del sistema ya que es la parte más importante en el desarrollo de software. El documento consta de cinco componentes primordiales que son: la vista de escenarios, vista lógica, vista física, vista de desarrollo y la vista de procesos, estos componentes nos ayudarán a tener una vista más clara acerca de las funcionalidades que el sistema va a poseer.

2. Propósito

El propósito de este documento es el de describir la arquitectura general del sistema con el uso del modelo 4+1, en el que se detalla los módulos del sistema, sus relaciones y el comportamiento general de cada uno de estos.

3. Referencias

Referencia	Título del Documento
Anexo 2(de la memoria del Trabajo de Titulación)	Especificación de Requerimientos de Software
Modelo	Arquitectura de software 4+1

4. Representación de la Arquitectura

La aplicación web tiene la arquitectura cliente-servidor, la misma que está construida por las vistas 4+1 de Kruchten y está basada en UML (Lenguaje Unificado de Modelado). En la siguiente tabla se describen las vistas del modelo.

Tabla 42: Vistas del Modelo 4 + 1

Vistas	UML	Descripción
Vista de Escenarios	Casos de Uso	Muestra la interacción entre los actores del sistema
Vista Lógica	Diagrama de Clases	Representa las funcionalidades y el servicio que proporciona a los usuarios

Vista Física	Diagrama de Despliegue	Muestra los componentes físicos de la solución informática
Vista de Desarrollo	Diagrama de Componentes	Detalla los componentes de la aplicación web para que así el programador visualice y comprenda las relaciones existentes y sus interacciones entre ellos.
Vista de Procesos	Diagrama de Actividad	Describe los procesos de cada una de las funcionalidades del sistema web.

5. Objetivos de la arquitectura

La aplicación web debe cumplir con las siguientes características.

- **Disponibilidad:** La aplicación web estará disponible 24/7 para efectos de la predicción del riesgo cardiovascular y el cálculo del perfil lipídico.
- **Rendimiento:** La aplicación web tendrá un rendimiento alto, ya que el framework django cuenta con su propia base de datos ya que las consultas realizadas a la base de datos serán más eficientes con un tiempo de respuesta bajo.
- **Portabilidad:** La aplicación web está construida en el Framework Django por lo que tiene una gran compatibilidad con navegadores y a su vez por ser una aplicación web se puede acceder desde cualquier sistema operativo mediante un navegador de internet.
- **Calidad:** La aplicación web debe cumplir con todos los requerimientos establecidos en la ERS.

6. Vista de Escenarios.

En esta primera sección se muestra el diagrama de casos de uso en el que se observa la interacción entre cada uno de los actores y la aplicación web. El siguiente diagrama presentado consta de 3 actores Administrador, Cobrador y Contribuyente, cada uno estos tienen diferentes funciones y para realizar estas deben estar autenticados correctamente en el sistema.

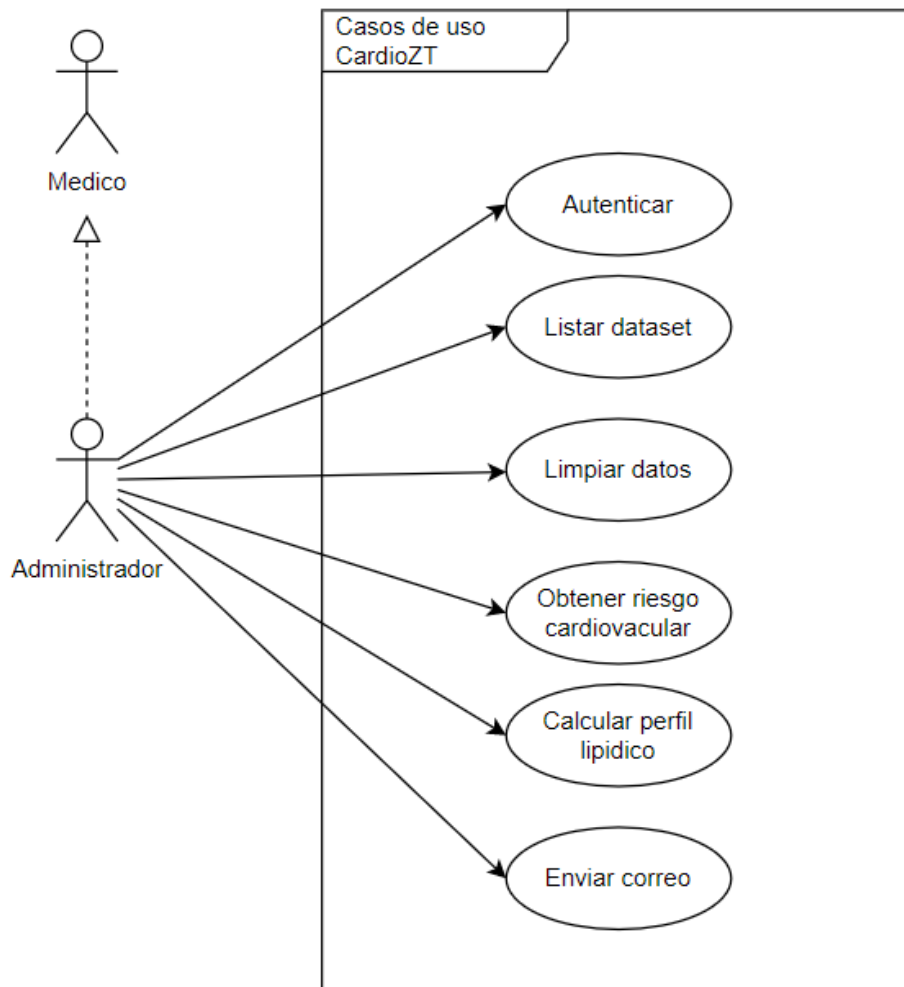


Figura 41: Diagrama de Casos de Uso (Fuente Propia)

7. Vista Lógica

En esta vista se refleja el diagrama de clases donde se exponen los requerimientos funcionales de la aplicación web.

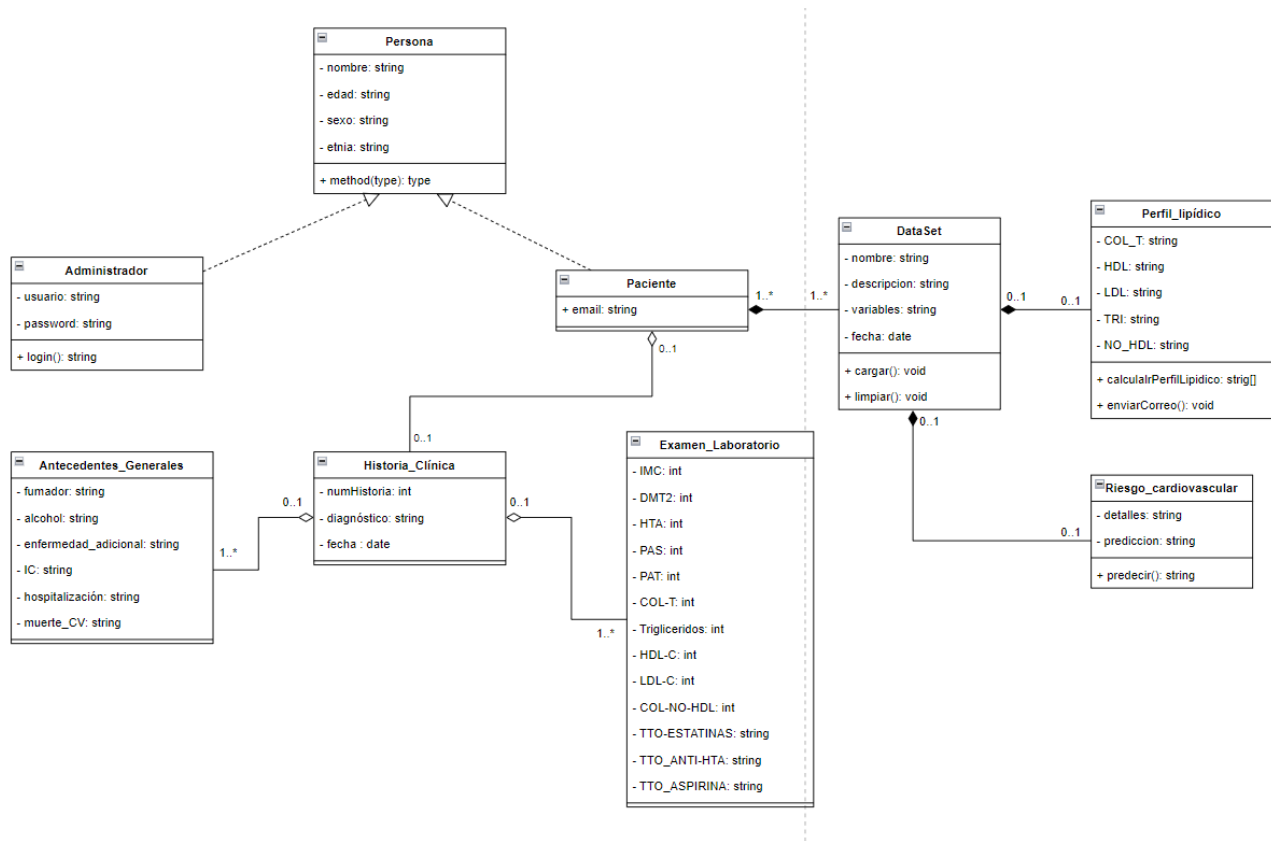


Figura 42: Diagrama de Clases (Fuente Propia)

8. Vista Física

En esta vista se muestra el diagrama de despliegue en el que se muestra cada uno de los nodos que conforma la aplicación web, utilizando la arquitectura cliente-servidor.

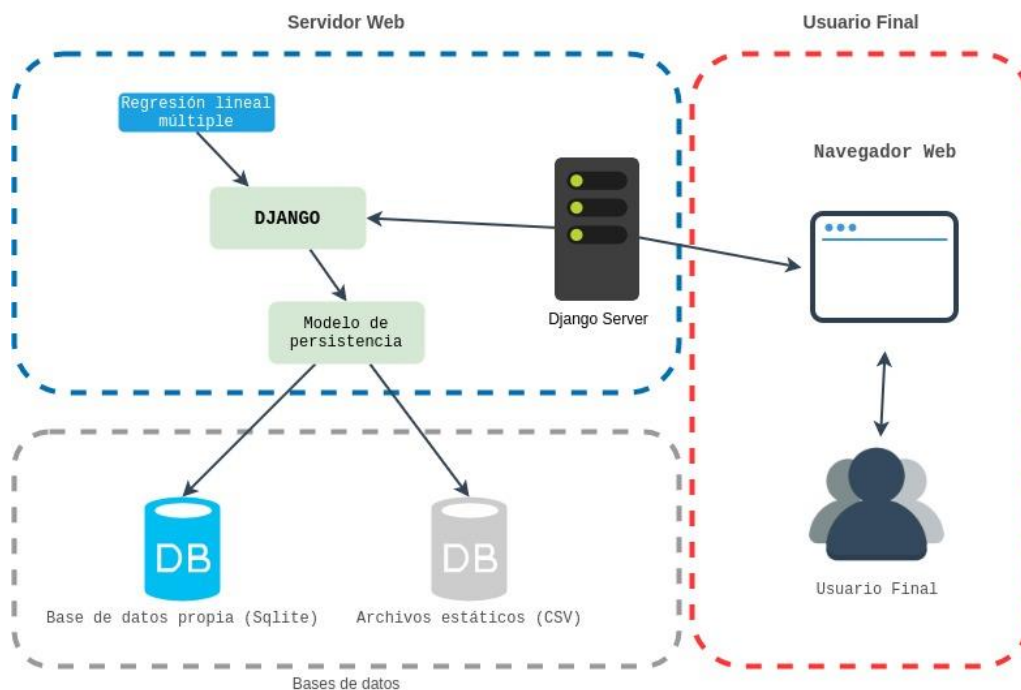


Figura 43: Diagrama de despliegue (Fuente Propia)

9. Vista de Desarrollo

En esta vista se muestra la forma de cómo están organizados los distintos componentes de la aplicación web en un ambiente de desarrollo de software.

- ✓ **Servidor web:** Se observa aquí el framework Django, que sigue la filosofía de baterías incluidas, en el cual se programa la lógica de negocio a través de los archivos estáticos csv cargados, activado el entorno virtualenv y levantado el servidor propio de django para las peticiones y mostradas a través de las vistas de acuerdo a cada template solicitado.
- ✓ **Aplicación Web:** Se observa aquí el aplicativo CardioZT, el cual consta de 4 módulos, el primero llamado Modelos en el cual el usuario puede ejecutar el modelo predictivo para el pronóstico del riesgo cardiovascular, el segundo de nombre Listar datos, el cual muestra al usuario la lista de datasets guardados. En el tercer módulo llamado Limpiar datos el usuario puede limpiar un dataset seleccionado con función automática. En último módulo de nombre Perfil lipídico el usuario insertar las variables necesarias para el cálculo matemático y registro de sus resultados.

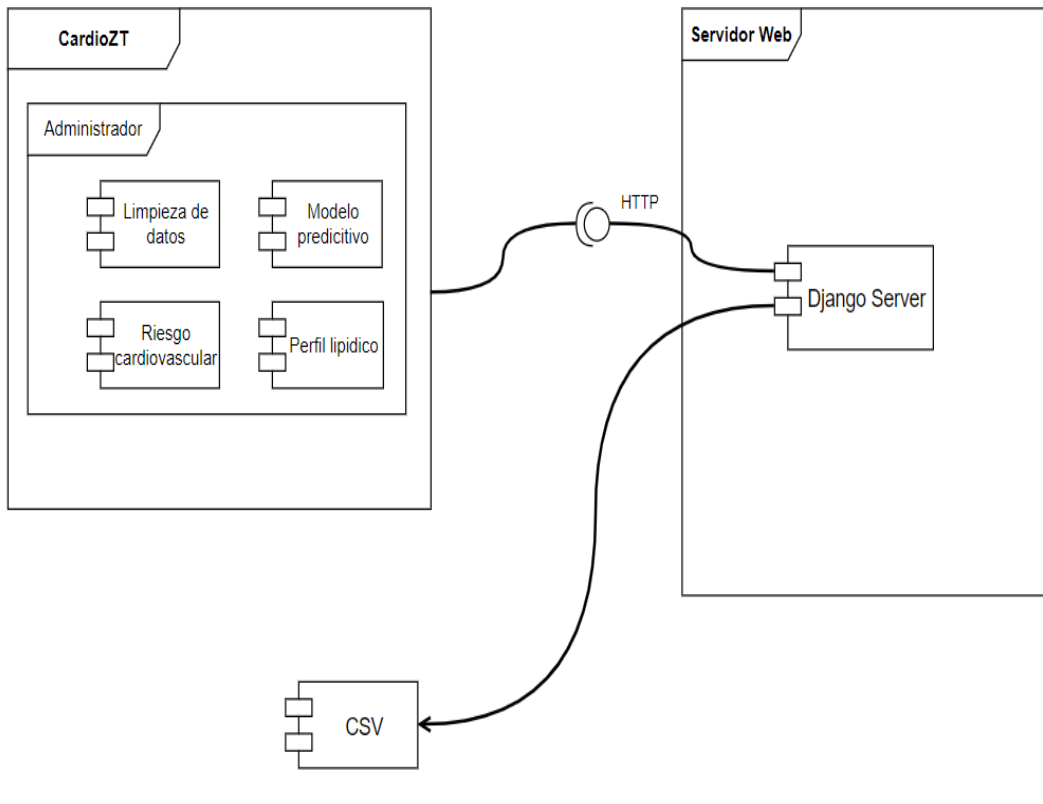


Figura 44: Diagrama de componentes (Fuente: Propia)

10. Vista de Procesos

En esta vista se presenta los diferentes diagramas de actividad donde se puede observar el flujo de un proceso que es realizado mediante la interacción del usuario y la aplicación web desde el inicio hasta su finalización.

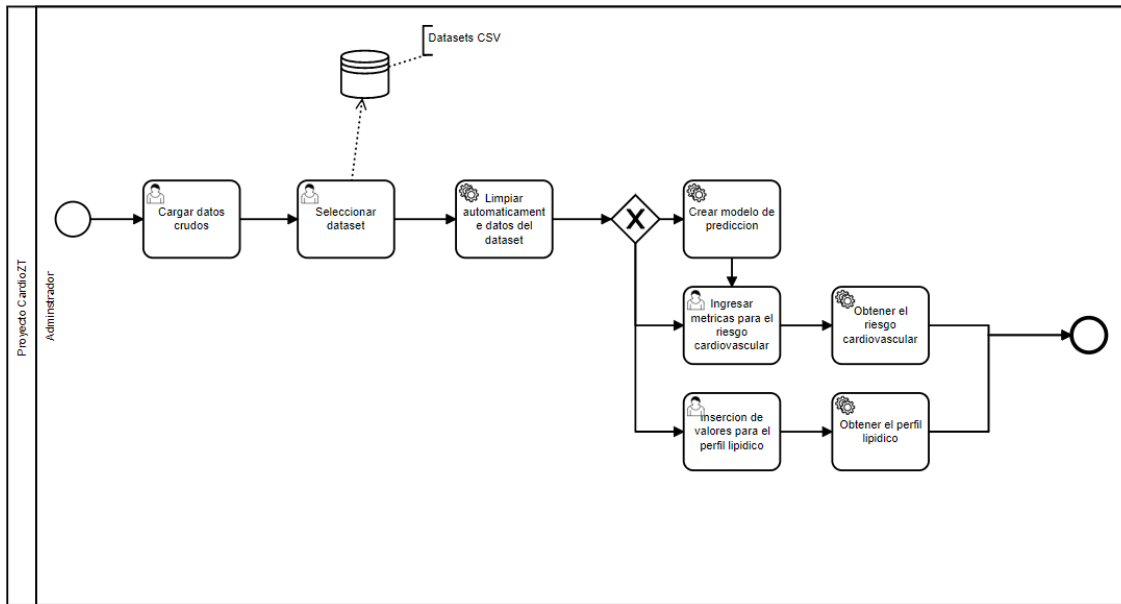


Figura 45: Diagrama de Actividad: Predicción y cálculo de pacientes con DMT2 y HTA

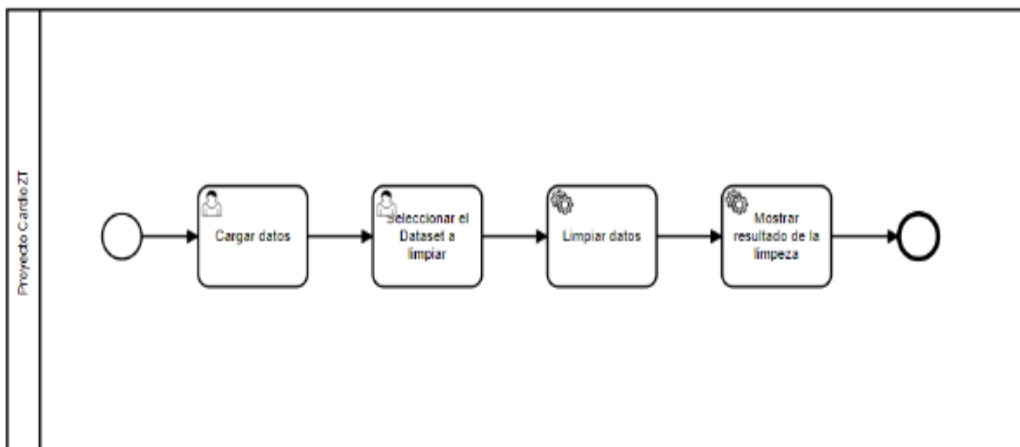


Figura 46: Diagrama de Actividad: Limpieza de conjunto total de datos de pacientes

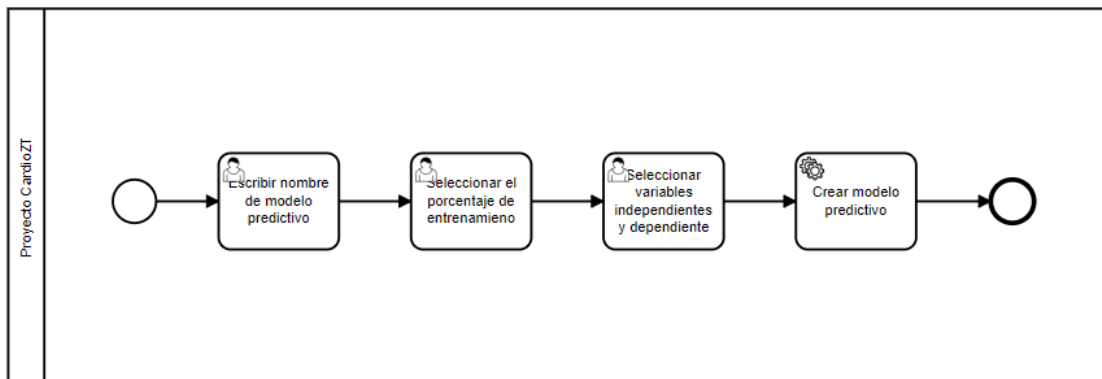


Figura 47: Diagrama de Actividad: Creación del modelo predictivo

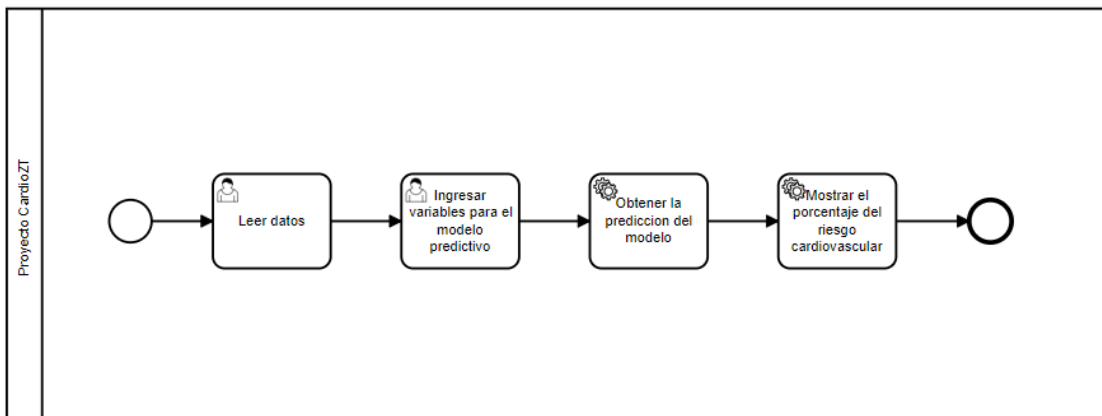


Figura 48: Diagrama de Actividad: Obtener el riesgo cardiovascular_Modelo Predictivo

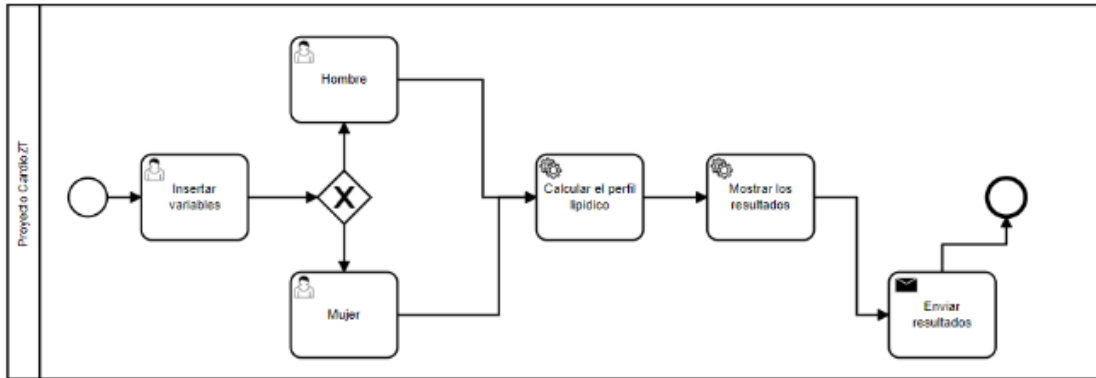


Figura 49: Diagrama de Actividad: Obtener el perfil lipídico_Modelo Matemático.

Anexo 7: Historias de Usuario.

Historias de usuario

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos

1. Introducción

En el presente documento se detalla las Historias de Usuario estas fueron obtenidas de los requerimientos funcionales del sistema y definidas conjuntamente con el usuario, esto es necesario para realizar la codificación del sistema mediante esprints y también para su posterior uso en las pruebas de aceptación.

2. Propósito

El propósito de este documento es el de definir conjuntamente con el usuario final las Historias de Usuario necesarias para dar inicio con la codificación del sistema, esto servirá al programador para establecer los tiempos de desarrollo de todas las funcionalidades y a su vez servirá para realizar las pruebas de aceptación

3. Referencias

Referencia	Título del Documento
Anexo 4(de la memoria del Trabajo de Titulación)	Especificación de Requerimientos de Software

4. Requerimientos del Sistema

Código	Requerimiento
RF01	Autenticar usuario
RF02	Cargar los datos cuantitativos de laboratorio del paciente
RF03	Limpiar los datos cuantitativos del conjunto de pacientes
RF04	Mostrar los datos normalizados en base a métricas requeridas del sistema
RF05	Ingresar datos para predecir el riesgo cardiovascular
RF06	Mostrar el porcentaje del riesgo cardiovascular del paciente
RF07	Mostrar gráficos estadísticos de los datos cargados en el sistema
RF08	Ingresar datos para calcular el perfil lipídico
RF09	Mostrar el nivel del perfil lipídico del paciente
RF10	Enviar mediante correo los resultados obtenidos del perfil lipídico

5. Definición de las Historias de Usuario

En esta sección se definen las historias de usuario correspondientes al sistema web para la administración y cobro del servicio de agua potable.

Tabla 43: Historia de Usuario: Autenticar usuario

Historia de Usuario	
Número: 1	Rol: Administrador
Nombre Historia: Autenticar usuario	
Prioridad en negocio: Media (Alta/Media/Baja)	Riesgo en desarrollo: Medio (Alto/Medio/Bajo)
Puntos estimados: 1	Iteración asignada: 1
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito autenticarme con el usuario y contraseña.	
Validación: RF01	
Criterios de Aceptación: Validar los campos de usuario y contraseña que no estén vacíos. Ingresar a la pantalla principal cuando el usuario sea autenticado correctamente. Mostrar mensaje de Bienvenida al sistema	

Tabla 44: Historia de Usuario: Cargar los datos cuantitativos de laboratorio del paciente

Historia de Usuario	
Número: 2	Rol: Administrador
Nombre Historia: Cargar los datos cuantitativos de laboratorio del paciente.	
Prioridad en negocio: Alta (Alta/Media/Baja)	Riesgo en desarrollo: Bajo (Alto/Medio/Bajo)
Puntos estimados: 1	Iteración asignada: 1
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito cargar los datos cuantitativos de laboratorio del paciente	
Validación: RF02	
Criterios de Aceptación: Cargar el dataset de los registros de los pacientes en una lista Verificar los nombres y fechas de los datasets.	

Tabla 45:Historia de Usuario: Limpiar los datos cuantitativos del conjunto total de pacientes

Historia de Usuario	
Número: 3	Rol: Administrador
Nombre Historia: Limpiar los datos cuantitativos del conjunto total de pacientes	
Prioridad en negocio: Media (Alta/Media/Baja)	Riesgo en desarrollo: Medio (Alto/Medio/Bajo)
Puntos estimados: 1.5	Iteración asignada: 1
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito limpiar el total de datos de los archivos csv	
Validación: RF03	
Criterios de Aceptación: Mostrar una ventana para seleccionar dataset a limpiar Ingresar nombre y una descripción del dataset Obtener datos limpios del dataset	

Tabla 46:Historia de Usuario: Mostrar los datos normalizados en base a métricas requeridas

Historia de Usuario	
Número: 4	Rol: Administrador
Nombre Historia: Mostrar los datos normalizados en base a métricas requeridas del sistema	
Prioridad en negocio: Media (Alta/Media/Baja)	Riesgo en desarrollo: Bajo (Alto/Medio/Bajo)
Puntos estimados: 1	Iteración asignada: 1
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito seleccionar las variables del dataset a utilizar para hacer el pronóstico del riesgo cardiovascular	
Validación: RF04	
Criterios de Aceptación: Mostrar dataset a analizar por nombres. Permitir seleccionar las variables para hacer la predicción del riesgo c.	

Tabla 47: Historia de Usuario: Ingresar datos para predecir el riesgo cardiovascular

Historia de Usuario	
Número: 5	Rol: Administrador
Nombre Historia: Ingresar datos para predecir el riesgo cardiovascular	
Prioridad en negocio: Alta (Alta/Media/Baja)	Riesgo en desarrollo: Medio (Alto/Medio/Bajo)
Puntos estimados: 2	Iteración asignada: 1
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito ingresar los valores para el riesgo cardiovascular	
Validación: RF05	
Criterios de Aceptación: Validar si están llenos los campos correctamente. Mensaje de advertencia de forma de ingresar valores a los campos de sexo, edad, dmt2, fumador y tto-anti-hta.	

Tabla 48: Historia de Usuario: Mostrar el porcentaje del riesgo cardiovascular del paciente

Historia de usuario	
Número: 6	Rol: Administrador
Nombre Historia: Mostrar el porcentaje del riesgo cardiovascular del paciente	
Prioridad en negocio: Alta (Alta/Media/Baja)	Riesgo en desarrollo: Medio (Alto/Medio/Bajo)
Puntos estimados: 1	Iteración asignada: 2
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito visualizar el resultado del riesgo c. del paciente	
Validación: RF06	
Criterios de Aceptación: Visualizar la estructura de la regresión lineal múltiple para hacer predicción. Leer la etiqueta: bajo, moderado, alto y muy alto riesgo.	

Tabla 49: Historia de Usuario: Mostrar gráficos estadísticos de los datos cargados en el sistema

Historia de Usuario	
Número: 7	Rol: Administrador
Nombre Historia: Mostrar gráficos estadísticos de los datos cargados en el sistema	
Prioridad en negocio: Baja (Alta/Media/Baja)	Riesgo en desarrollo: Bajo (Alto/Medio/Bajo)
Puntos estimados: 1	Iteración asignada: 2
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito ver las estadísticas de los datos cargados en el sistema	
Validación: RF07	
Criterios de Aceptación: Mostrar el género y edad del conjunto total de registros. Mostrar el total de pacientes diabéticos e hipertensos. Mostrar el porcentaje de tratamiento antihipertensivo.	

Tabla 50: Historia de Usuario: Ingresar datos para calcular el perfil lipídico

Historia de Usuario	
Número: 8	Rol: Administrador
Nombre Historia: Ingresar datos para calcular el perfil lipídico	
Prioridad en negocio: Media (Alta/Media/Baja)	Riesgo en desarrollo: Medio (Alto/Medio/Bajo)
Puntos estimados: 1	Iteración asignada: 2
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito calcular el perfil lipídico del paciente	
Validación: RF08	
Criterios de Aceptación: Validar los campos de texto requeridos al momento de ingresar los datos Permita seleccionar género: hombre y mujer.	

Tabla 51: Historia de Usuario: Mostrar el nivel del perfil lipídico del paciente

Historia de Usuario	
Número: 9	Rol: Administrador
Nombre Historia: Mostrar el nivel del perfil lipídico del paciente.	
Prioridad en negocio: Alta (Alta/Media/Baja)	Riesgo en desarrollo: Alta (Alto/Medio/Bajo)
Puntos estimados: 2	Iteración asignada: 2
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema requiero ver el nivel del perfil lipídico que maneja el paciente en prevención secundaria.	
Validación: RF09	
Criterios de Aceptación: Obtener el estado del perfil lipídico del paciente. Permita visualizar el cálculo del perfil lipídico en banda de colores para sus resultados	

Tabla 52: Historia de Usuario: Enviar por correo los resultados obtenidos del perfil lipídico

Historia de Usuario	
Número: 10	Rol: Administrador
Nombre Historia: Enviar por correo los resultados obtenidos del perfil lipídico	
Prioridad en negocio: Media (Alta/Media/Baja)	Riesgo en desarrollo: Media (Alto/Medio/Bajo)
Puntos estimados: 1	Iteración asignada: 3
Programador responsable: Roger Torres	
Descripción: Como médico administrador del sistema necesito enviar por correo electrónico el resultado del nivel del perfil lipídico del paciente.	
Validación: RF10	
Criterios de Aceptación: Mostrar los niveles del perfil lipídico en una caja. Permitir enviar el resultado al correo electrónico: aplicación.perfil@gmail.com Aviso de acción realizada satisfactoriamente.	

6. Resumen de las Historias de Usuario

En esta sección se presenta un resumen de las historias de usuario definidas anteriormente para un mejor entendimiento.

Tabla 53: Resumen de las Historias de Usuario

N°	Nombre	Prioridad	Riesgo	Iteración
1	Autenticar usuario.	Alta	Medio	1
2	Cargar los datos cuantitativos de laboratorio del paciente.	Alta	Bajo	1
3	Limpiar los datos cuantitativos del conjunto total de pacientes	Media	Medio	1
4	Mostrar los datos normalizados en base a métricas requeridas del sistema	Media	Bajo	1
5	Ingresar datos para predecir el riesgo cardiovascular	Alta	Medio	1
6	Mostrar el porcentaje del riesgo cardiovascular del paciente	Alta	Medio	2
7	Mostrar gráficos estadísticos de los datos cargados en el sistema	Baja	Bajo	2
8	Ingresar datos para calcular el perfil lipídico	Media	Medio	2
9	Mostrar el nivel del perfil lipídico del paciente	Alta	Alto	2
10	Enviar por correo los resultados del perfil lipídico.	Media	Medio	3

Anexo 8: Prototipo de Interfaz.

Pantallas de aceptación del módulo de software

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos

1. Introducción

En este apartado se presentan los prototipos de pantalla generados durante la codificación del sistema como una parte justificativa de la revisión técnica, del mismo modo estos sirvieron para la validación visual de los requerimientos del software. No obstante, si existe algún cambio en las interfaces, ésta se lo realizara bajo un control de cambios y previa justificación.

2. Interfaces del Sistema

En esta sección del documento se presentan todas las interfaces que se desarrollaron para el sistema CardioZT y a la misma ves se presenta la descripción de cada una de ellas.

Tabla 54: Pantalla de inicio de sesión

Pantalla de Inicio de Sesión CardioZT		
Numero de pantalla	001	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Mostrar al administrador o contribuyente que datos debe ingresar para realizar un ingreso exitoso.	

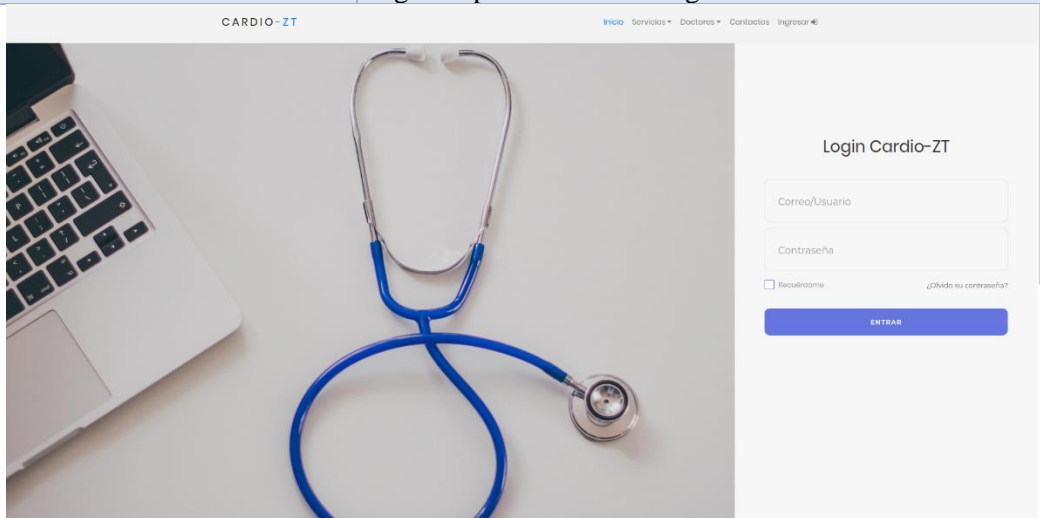
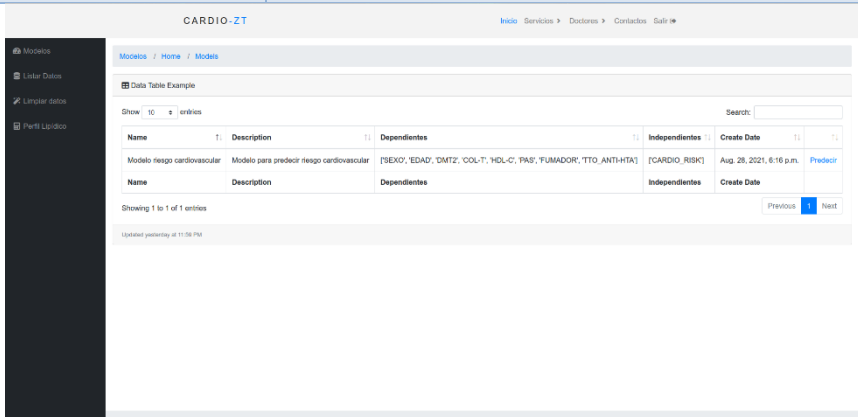


Tabla 55: Pantalla principal del médico administrador

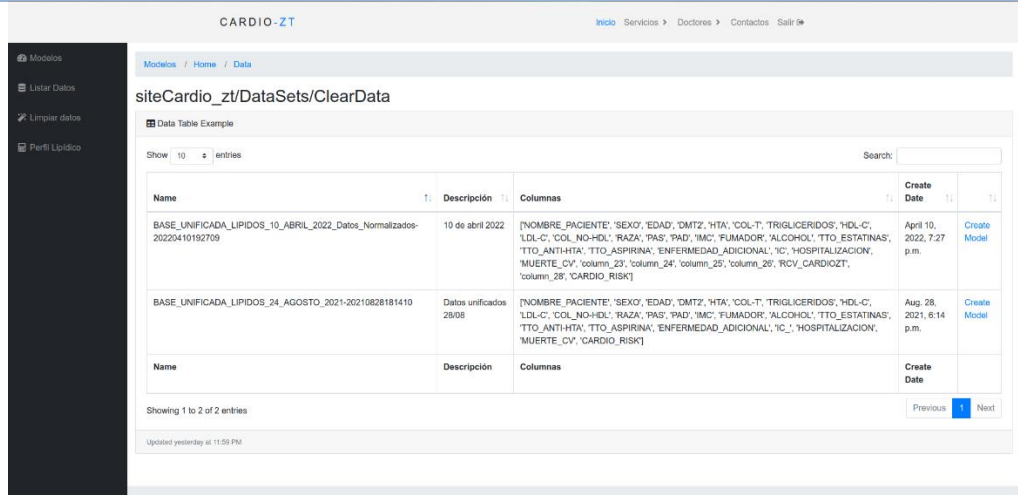
Pantalla principal del Médico Administrador		
Numero de pantalla	002	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Presentar al usuario información general de todo lo referente al perfil lipídico y su relación con el riesgo cardiovascular.	



The screenshot shows a web application interface for 'CARDIO-ZT'. It features a dark sidebar with navigation options: 'Modelos', 'Listar Datos', 'Limpiar datos', and 'Perfil Lipídico'. The main content area displays a table titled 'Data Table Example' with columns: Name, Description, Dependientes, Independientes, and Create Date. A single entry is visible: 'Modelo riesgo cardiovascular' with a description 'Modelo para predecir riesgo cardiovascular' and a 'Create Date' of 'Aug 28, 2021, 6:16 p.m.'. A 'Predecir' button is located to the right of the entry. The interface includes pagination controls and a search bar.

Tabla 56: Pantalla principal del listado de datos

Pantalla principal del listado de datos		
Numero de pantalla	003	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Mostrar el listado de los datasets cargados en el sistema para el riesgo cardiovascular.	



The screenshot shows the 'CARDIO-ZT' application interface for listing data sets. The sidebar includes 'Modelos', 'Listar Datos', 'Limpiar datos', and 'Perfil Lipídico'. The main content area is titled 'siteCardio_zt/DataSets/ClearData' and shows a table with columns: Name, Descripción, Columnas, and Create Date. Two entries are listed:

- Entry 1:** Name: 'BASE_UNIFICADA_LIPIDOS_10_ABRIL_2022_Datos_Normalizados-20220410192709', Descripción: '10 de abril 2022', Columnas: '[NOMBRE_PACIENTE, 'SEXO', 'EDAD', 'DMTZ', 'HTA', 'COL-T', 'TRIGLICERIDOS', 'HDL-C', 'LDL-C', 'COL_NO-HDL', 'RAZA', 'PAS', 'PAD', 'IMC', 'FUMADOR', 'ALCOHOL', 'TTO_ESTATINAS', 'TTO_ANTIHTA', 'TTO_ASPIRINA', 'ENFERMEDAD_ADICIONAL', 'IC', 'HOSPITALIZACION', 'MUERTE_CV', 'column_23', 'column_24', 'column_25', 'column_26', 'RCV_CARDIOZT', 'column_26', 'CARDIO_RISK']', Create Date: 'April 10, 2022, 7:27 p.m.'.
- Entry 2:** Name: 'BASE_UNIFICADA_LIPIDOS_24_AGOSTO_2021-20210828181410', Descripción: 'Datos unificados 28/08', Columnas: '[NOMBRE_PACIENTE, 'SEXO', 'EDAD', 'DMTZ', 'HTA', 'COL-T', 'TRIGLICERIDOS', 'HDL-C', 'LDL-C', 'COL_NO-HDL', 'RAZA', 'PAS', 'PAD', 'IMC', 'FUMADOR', 'ALCOHOL', 'TTO_ESTATINAS', 'TTO_ANTIHTA', 'TTO_ASPIRINA', 'ENFERMEDAD_ADICIONAL', 'IC', 'HOSPITALIZACION', 'MUERTE_CV', 'CARDIO_RISK']', Create Date: 'Aug 28, 2021, 6:14 p.m.'.

 The interface includes pagination controls and a search bar.

Tabla 57: Pantalla de administración de limpieza de datos

Pantalla de administración de limpieza de datos		
Numero de pantalla	004	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Presentar la ventana de limpieza de datos con función automática.	

Tabla 58: Pantalla de creación de modelo automático

Pantalla de creación del Modelo Automático		
Numero de pantalla	005	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Visualizar los elementos para poder hacer la predicción y el porcentaje de entrenamiento.	

Tabla 59: Pantalla principal de riesgo cardiovascular

Pantalla principal del riesgo cardiovascular		
Numero de pantalla	006	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Ventana principal para hacer la predicción del riesgo cardiovascular del paciente seleccionado del conjunto de datos	

Tabla 60: Pantalla principal de obtención del riesgo cardiovascular

Pantalla de Obtención del riesgo cardiovascular		
Numero de pantalla	007	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Mostrar el porcentaje de la predicción del riesgo cardiovascular del paciente y su respectiva etiqueta de categoría del mismo.	

Tabla 61: Pantalla de estadísticas de los pacientes

Pantalla de Estadísticas de los Pacientes		
Numero de pantalla	008	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Mostrar los porcentajes del conjunto total de datos de los pacientes en género, deceso y rangos de edad con el modelo predictivo.	
	<p>The figure displays three pie charts. The first chart, titled 'Total de pacientes del conjunto de datos', shows the distribution of patients by gender (Hombre in blue, Mujeres in red). The second chart, 'Personas fallecidas basado en genero', shows the distribution of deceased patients by gender. The third chart, 'Rango de edades', shows the distribution of patients by age range (1-17 in green, 18-64 in blue, 65-110 in red).</p>	

Tabla 62: Pantalla de estadísticas de las enfermedades asociadas

Pantalla de Estadísticas de las Enfermedades Asociadas		
Numero de pantalla	009	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Mostrar los porcentajes del conjunto total de datos de la enfermedad de diabetes tipo 2 e hipertensión arterial sistémica con el modelo predictivo.	
	<p>The figure is a stacked bar chart showing the number of patients for three associated diseases: HTA, DM2, and T10-ANTI-HTA. The y-axis represents the number of patients, ranging from 0 to 6000. The x-axis lists the diseases. Each bar is divided into two categories: SI (green) and NO (pink). The total number of patients for each disease is approximately 5500.</p>	

Tabla 63: Pantalla de inserción de variables para el perfil lipídico

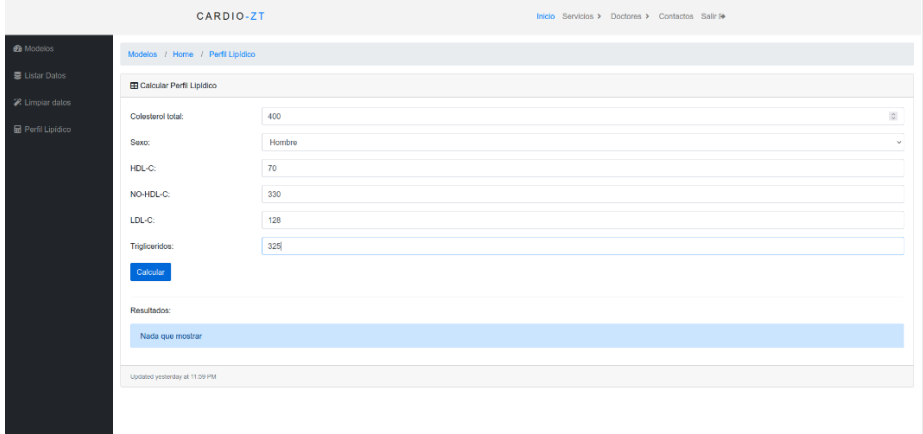
Pantalla de inserción de variables para el perfil lipídico		
Numero de pantalla	010	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Presentar ventana para la inserción de variables para el cálculo del nivel del perfil lipídico del paciente.	
		

Tabla 64: Pantalla del nivel del perfil lipídico

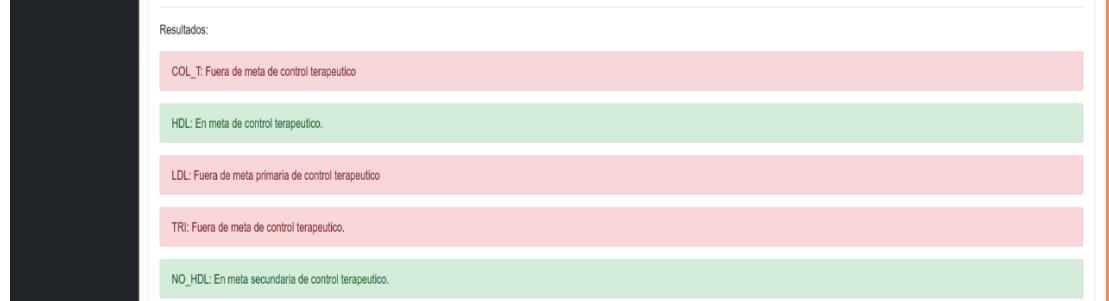
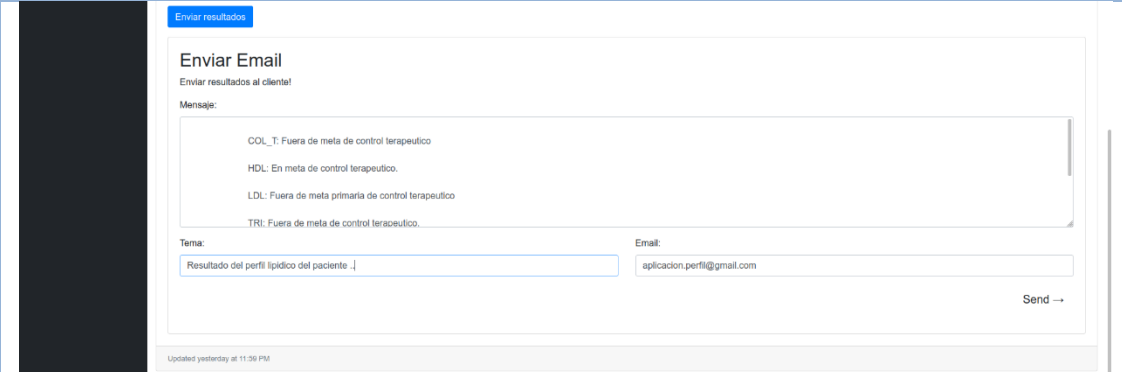
Pantalla del nivel del perfil lipídico		
Numero de pantalla	011	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Listar el nivel de perfil lipídico del paciente de acuerdo a las variables establecidas en el estudio.	
		

Tabla 65: Pantalla de envío de resultados del perfil lipídico

Pantalla de envío de resultados del perfil lipídico		
Numero de pantalla	012	
Estado de Prueba	Aceptado sin cambios	X
	Aceptado con cambios	
	Rediseñar	
Objetivo de la plantilla	Mostrar al usuario el envío del resultado del nivel del perfil lipídico para guardar la información requerida por el paciente diabético y/o hipertenso.	
 <p>The screenshot shows a web interface for sending an email. At the top left, there is a blue button labeled 'Enviar resultados'. Below it, the title 'Enviar Email' is displayed, followed by the subtitle 'Enviar resultados al cliente!'. A 'Mensaje:' label is positioned above a large text area containing the following text: 'COL_T: Fuera de meta de control terapeutico', 'HDL: En meta de control terapeutico.', 'LDL: Fuera de meta primaria de control terapeutico', and 'TRI: Fuera de meta de control terapeutico.'. Below the message area, there are two input fields: 'Tema:' with the value 'Resultado del perfil lipidico del paciente .j' and 'Email:' with the value 'aplicacion.perfil@gmail.com'. A 'Send -->' button is located at the bottom right of the form. At the very bottom of the page, a small text reads 'Updated yesterday at 11:56 PM'.</p>		

Anexo 9: Pruebas de Aceptación.

Pruebas de Aceptación

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos

1. Introducción

El presente documento detalla las pruebas de Aceptación de la aplicación web para determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos del Hospital Ceibos de Guayaquil, a partir de la socialización del sistema con el personal de salud encargado de cuidado y control de los pacientes, quienes validan la funcionalidad y los requerimientos establecidos de la aplicación.

2. Propósito

El propósito de esto es validar la funcionalidad de cada uno de los módulos y verificar el cumplimiento de los requerimientos establecidos en la aplicación web mediante la socialización del sistema con el personal de salud encargado de cuidado y control de los pacientes, quienes validan la funcionalidad y los requerimientos establecidos de la aplicación, y con esto comprobar que la aplicación web cuenta con su alto nivel de aceptación.

3. Referencias

Referencia	Título del Documento
Anexo 4(de la memoria del Trabajo de Titulación)	Especificación de Requerimientos de Software

4. Requerimientos del Sistema

Tabla 66: Requerimientos Funcionales

Código	Requerimiento
RF01	Autenticar usuario
RF02	Cargar los datos cuantitativos de laboratorio del paciente
RF03	Limpiar los datos cuantitativos del conjunto total de pacientes
RF04	Mostrar los datos normalizados en base a métricas requeridas
RF05	Ingresar datos para predecir el riesgo cardiovascular
RF06	Mostrar el porcentaje del riesgo cardiovascular del paciente
RF07	Mostrar gráficos estadísticos de los datos cargados en el sistema
RF08	Ingresar datos para calcular el perfil lipídico
RF09	Mostrar el nivel del perfil lipídico del paciente
RF10	Enviar mediante correo los resultados obtenidos del perfil lipídico

5. Casos de Prueba

Tabla 67: Caso de prueba nro. 1

CASO DE PRUEBA	
Historia de Usuario: Autenticar usuario	Nro. Historia de Usuario: 1
Condiciones de Ejecución: Como médico administrador del sistema necesito autenticarme con el usuario y contraseña.	
Criterios de Validación: Validar los campos de usuario y contraseña que no estén vacíos. Ingresar a la pantalla principal cuando el usuario sea autenticado correctamente. Mostrar mensaje de Bienvenida al sistema	
Resultado Esperado: Inicio de sesión correcto	
Validado: Exitosamente	

Tabla 68: Caso de prueba nro. 2

CASO DE PRUEBA	
Historia de Usuario: Cargar los datos cuantitativos de Laboratorio del paciente	Nro. Historia de Usuario: 2
Condiciones de Ejecución: Como médico administrador del sistema necesito cargar los datos cuantitativos de laboratorio del paciente	

<p>Criterios de Validación:</p> <p>Cargar el dataset de los registros de los pacientes en una lista</p> <p>Verificar los nombres y fechas de los datasets.</p>
<p>Resultado Esperado: Carga de datos de pacientes correctamente</p>
<p>Validado: Exitosamente</p>

Tabla 69: Caso de prueba nro. 3

CASO DE PRUEBA	
<p>Historia de Usuario: Limpiar los datos cuantitativos del conjunto total de pacientes</p>	<p>Nro. Historia de Usuario: 3</p>
<p>Condiciones de Ejecución: Como médico administrador del sistema necesito limpiar el total de datos de los archivos csv.</p>	
<p>Criterios de Validación:</p> <p>Mostrar una ventana para seleccionar dataset a limpiar</p> <p>Ingresar nombre y una descripción del dataset</p> <p>Obtener datos limpios del dataset</p>	
<p>Resultado Esperado: Datos de pacientes limpios correctamente</p>	
<p>Validado: Exitosamente</p>	

Tabla 70: Caso de prueba nro. 4

CASO DE PRUEBA	
Historia de Usuario: Mostrar los datos normalizados en base a métricas requeridas	Nro. Historia de Usuario: 4
Condiciones de Ejecución: Como médico administrador del sistema necesito seleccionar las variables del dataset a utilizar para hacer el pronóstico del riesgo cardiovascular.	
Criterios de Validación: Mostrar dataset a analizar por nombres. Permitir seleccionar las variables para hacer la predicción del riesgo cardiovascular.	
Resultado Esperado: Los datos están normalizados correctamente.	
Validado: Exitosamente	

Tabla 71: Caso de prueba nro. 5

CASO DE PRUEBA	
Historia de Usuario: Ingresar datos para predecir el riesgo cardiovascular	Nro. Historia de Usuario: 5
Condiciones de Ejecución: Como médico administrador del sistema necesito ingresar los valores para el riesgo cardiovascular	
Criterios de Validación: Validar si están llenos los campos correctamente.	

Mensaje de advertencia de forma de ingresar valores a los campos de sexo, edad, dmt2, fumador y tto-anti-hta
Resultado Esperado: Campos de variables para el riesgo ingresados correctamente.
Validado: Exitosamente

Tabla 72: Caso de prueba nro.6

CASO DE PRUEBA	
Historia de Usuario: Mostrar el porcentaje del riesgo cardiovascular del paciente	Nro. Historia de Usuario: 6
Condiciones de Ejecución: Como médico administrador del sistema necesito visualizar el resultado del riesgo c. del paciente	
Criterios de Validación: Visualizar la estructura de la regresión lineal múltiple para hacer predicción. Leer la etiqueta: bajo, moderado, alto y muy alto riesgo.	
Resultado Esperado: Resultado correcto del riesgo cardiovascular.	
Validado: Exitosamente	

Tabla 73: Caso de prueba nro. 7

CASO DE PRUEBA	
Historia de Usuario: Mostrar gráficos estadísticos de los datos cargados en el sistema	Nro. Historia de Usuario: 7

<p>Condiciones de Ejecución: Como médico administrador del sistema necesito ver las estadísticas de los datos cargados en el sistema.</p>
<p>Criterios de Validación:</p> <p>Mostrar el género y edad del conjunto total de registros.</p> <p>Mostrar el total de pacientes diabéticos e hipertensos.</p> <p>Mostrar el porcentaje de tratamiento antihipertensivo.</p>
<p>Resultado Esperado: Estadísticas correcta del conjunto total de pacientes.</p>
<p>Validado: Exitosamente</p>

Tabla 74: Caso de prueba nro. 8

CASO DE PRUEBA	
<p>Historia de Usuario: Ingresar datos para calcular el perfil lipídico</p>	<p>Nro. Historia de Usuario: 8</p>
<p>Condiciones de Ejecución: Como médico administrador del sistema necesito calcular el perfil lipídico del paciente.</p>	
<p>Criterios de Validación:</p> <p>Validar los campos de texto requeridos al momento de ingresar los datos</p> <p>Permita seleccionar género: hombre y mujer.</p>	
<p>Resultado Esperado: Campos llenados correctamente y selección de género correcto.</p>	
<p>Validado: Exitosamente</p>	

Tabla 75: Caso de prueba nro. 9

CASO DE PRUEBA	
Historia de Usuario: Mostrar el nivel del perfil lipídico del paciente	Nro. Historia de Usuario: 9
Condiciones de Ejecución: Como médico administrador del sistema requiero ver el nivel del perfil lipídico que maneja el paciente en prevención secundaria	
Criterios de Validación: Obtener el estado del perfil lipídico del paciente. Permita visualizar el cálculo del perfil lipídico en banda de colores para sus resultados	
Resultado Esperado: Niveles de perfil lipídico del paciente correctos y banda de coloración exitosa.	
Validado: Exitosamente	

Tabla 76: Caso de prueba nro. 10

CASO DE PRUEBA	
Historia de Usuario: Enviar mediante correo los resultados obtenidos del perfil lipídico	Nro. Historia de Usuario: 10
Condiciones de Ejecución: Como médico administrador del sistema necesito enviar por correo electrónico el resultado del nivel del perfil lipídico del paciente.	
Criterios de Validación: Mostrar los niveles del perfil lipídico en una caja.	

Permitir enviar el resultado al correo electrónico: aplicación.perfil@gmail.com

Aviso de acción realizada satisfactoriamente.

Resultado Esperado: Envío correcto de resultados del perfil lipídico por correo.

Validado: Exitosamente



UNIVERSIDAD NACIONAL DE LOJA

FACULTAD DE LA ENERGÍA, LAS INDUSTRIAS Y LOS RECURSOS NATURALES NO RENOVABLES

Con la consecuente aprobación de las Pruebas de Aceptación de la Aplicación web para determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la consulta externa del Hospital Ceibos. Se deja constancia de la Validación de la herramienta empleada.

Nombre del Desarrollador: Roger Alexander Torres Yaguana

Para constancia firman:


Dr. Lenin Francisco Torres Yaguana
MEDICINA INTERNA
1006-2016-1752577 C.I. 1103504690
Instituto Ecuatoriano De Seguridad Social
HOSPITAL GENERAL DEL NORTE DE QUILLOA, LOS CEIBOS

.....
Dr. Lenin Torres Y.


Dr. Antonio Javier Pozo Garcia
MEDICINA INTERNA
1006-2016-1752578 C.I. 0920915758
Instituto Ecuatoriano De Seguridad Social
HOSPITAL GENERAL DEL NORTE DE QUILLOA, LOS CEIBOS
.....
Dr. Antonio Pozo G.

Anexo 10: Acta de Acuerdo de Confidencialidad.

Acta de Acuerdo de Confidencialidad

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos



INSTITUTO ECUATORIANO DE SEGURIDAD SOCIAL

COMPROMISO DE CONFIDENCIALIDAD DE LA INFORMACIÓN

Comparece **Roger Alexander Torres Yaguana**, portador de la cédula de ciudadanía número 1103491435 en mi calidad de **Estudiante de la Carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja** quien en adelante y para efectos del presente instrumento se denominará **“Estudiante en proceso de titulación”**, sujetándose a los siguientes términos y condiciones:

CLÁUSULA PRIMERA. - ANTECEDENTE:

- 1.1 El presente Compromiso de Confidencialidad se sustenta en la siguiente normativa: Constitución de la República del Ecuador (Art. 18 numeral 2, Art. 66 numerales 11 y 19, Art. 91, Art. 362); Ley Orgánica de Transparencia y Acceso a la Información Pública (Art. 6); Ley Orgánica de Salud (Art. 7 letra f), Ley Orgánica de Donación y Trasplante de Órganos, Tejidos y Células (Art. 11); Ley Orgánica del Servicio Público (Art. 22 letra j); Código Orgánico Administrativo (Art. 24); Código Orgánico Integral Penal (Art. 179 y Art. 229); Ley Orgánica del Sistema Nacional de Datos Públicos (Art. 4), Ley de Seguridad Social (Art. 247 inciso tercero), Ley de Derechos y Amparo del Paciente (Art. 4), Ley de Comercio Electrónico, Firmas Electrónicas y Mensajes de Datos (Art. 9), Normas de Control Interno para las Entidades, Organismos del Sector Público y Personas Jurídicas de Derecho Privado que dispongan de Recursos Públicos (400 Actividades de Control y 500-01 Controles sobre sistemas de información); Código de Ética del Instituto Ecuatoriano de Seguridad Social (Art. 9 y Art. 14 letras c y d).
- 1.2 El **“Estudiante en proceso de titulación”** y con previo aval del **Ing. Pablo Fernando Ordoñez Ordoñez-GESTOR ACADÉMICO DE LA CARRERA DE SISTEMAS Y COMPUTACIÓN DE LA UNIVERSIDAD NACIONAL DE LOJA**, conforme consta del Of. N° 100 CIS-FEIRNNR-UNL que se adjunta al presente instrumento como documento habilitante, requiere del **HOSPITAL GENERAL DEL NORTE DE GUAYAQUIL “LOS CEIBOS”**, se pueda obtener información relacionada con pacientes con Diabetes Mellitus Tipo 2 e Hipertensión Arterial Sistémica en el periodo comprendido desde mayo del 2017 a septiembre del 2020, información que servirá para el desarrollo de su trabajo de Titulación **“APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS”**.

CLÁUSULA SEGUNDA. - OBJETO:

- 2.1 El presente Compromiso de Confidencialidad tiene por finalidad proteger la información de propiedad institucional y de carácter personal de los pacientes del **HOSPITAL GENERAL DEL**



INSTITUTO ECUATORIANO DE SEGURIDAD SOCIAL

NORTE DE GUAYAQUIL “LOS CEIBOS”, y en general del asegurados y empleadores del Instituto Ecuatoriano de Seguridad Social, de los cuales cuya información se encuentra bajo su custodia en archivos físicos, bases de datos o almacenada en los recursos tecnológicos: sea impresa, digital o electrónica; y, aquella que se encuentre en etapa de gestión en procesos internos estratégicos, operativos o de apoyo, considerada como activo del **HOSPITAL GENERAL DEL NORTE DE GUAYAQUIL “LOS CEIBOS”** y del **IESS**, información catalogada como fundamental para el cumplimiento de su misión institucional, garantizando su disponibilidad, confidencialidad e integridad por parte de todos sus Directivos, Servidores y Terceros, en cumplimiento de la normativa legal vigente.

2.2 La información que se brindará de parte del **HOSPITAL GENERAL DEL NORTE DE GUAYAQUIL “LOS CEIBOS”**, al “Estudiante en proceso de titulación”, será la comprendida en el **periodo mayo del 2017 a septiembre del 2020**, y se enmarcará en forma estricta al pedido del **GESTOR ACADÉMICO DE LA CARRERA DE SISTEMAS DE LA UNIVERSIDAD NACIONAL DE LOJA** conforme consta del Of. N° 100 CIS-FEIRNNR-UNL y para su uso exclusivo en el desarrollo de su trabajo de Titulación “**APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS**” de autoría del “Estudiante en proceso de titulación”.

CLÁUSULA TERCERA. - COMPROMISO DEL ESTUDIANTE EN PROCESO DE TITULACIÓN:

El Estudiante en proceso de titulación se compromete con ética y profesionalismo de conformidad con las disposiciones legales vigentes a:



INSTITUTO ECUATORIANO DE SEGURIDAD SOCIAL

1. Guardar la confidencialidad y/o reserva de la información de propiedad institucional y la de carácter personal de los pacientes del Hospital General del Norte de Guayaquil, asegurados y empleadores del Instituto Ecuatoriano de Seguridad Social, que reciba, conozca, acceda, maneje o haga uso para el cumplimiento de su trabajo de titulación.
2. Utilizar en forma exclusiva y para fines académicos del trabajo de titulación antes referido, la información institucional y de carácter personal de los pacientes del Hospital General del Norte de Guayaquil, asegurados y empleadores del Instituto Ecuatoriano de Seguridad Social, garantizando en forma prolija la protección, disponibilidad, confidencialidad e integridad de la información, ya sea impresa, digital o electrónica, que se almacene en archivos físicos y bases de datos institucionales;
3. No establecer referencias de carácter personal de la información en el desarrollo del trabajo de titulación, con el fin de garantizar el derecho a la confidencialidad e intimidad de los pacientes, especialmente en datos inherentes a historia clínica y exámenes complementarios, de los cuales solo se autoriza el uso en cifras estadísticas y/o generales que serán empleadas en el trabajo de titulación;
4. Realizar la entrega-recepción y devolver al Hospital General del Norte de Guayaquil “Los Ceibos” o al IESS toda la información recibida y generada de forma física, digital o electrónica, a la finalización de las actividades referentes al desarrollo del trabajo de titulación.

CLÁUSULA CUARTA. - PROHIBICIONES:

El Estudiante en proceso de titulación queda prohibido a:

- a) Divulgar, distribuir, reproducir, utilizar, disponer, publicar por cualquier medio y/o para fines diferentes a los estipulados en el presente compromiso de confidencialidad, la información de propiedad institucional, carácter personal de los pacientes del Hospital General del Norte de Guayaquil “Los Ceibos”, asegurados y empleadores del IESS o de otras entidades públicas, que reciba, conozca, acceda, maneje o haga uso.
- b) Manipular la información de propiedad institucional y la de carácter personal de los pacientes del Hospital General del Norte de Guayaquil “Los Ceibos”, asegurados y empleadores del IESS, ya sea para beneficio propio o de terceros, o utilizarlas para propósitos diferentes, en trabajos presentes y futuros, que no sea el trabajo de titulación antes referido.
- c) Acceder a los sistemas y a las herramientas tecnológicas que registren, procesen, transmitan y almacenen la información de propiedad institucional y la de carácter personal de los pacientes del Hospital General del Norte de Guayaquil “Los Ceibos”, asegurados y empleadores del IESS, sin la supervisión de las autoridades y funcionarios de la institución.



INSTITUTO ECUATORIANO DE SEGURIDAD SOCIAL

- d) Utilizar los recursos tecnológicos del Hospital General del Norte de Guayaquil “Los Ceibos” o del IESS, como un medio de participación, acceso y distribución de actividades o materiales que vayan en contra de la normativa legal vigente o pongan en riesgo la confidencialidad e integridad de la información que se administra y custodia la institución; y,
- e) Perder, destruir o manipular intencionadamente o no los activos de información de la institución durante el tiempo de realización de las actividades inherentes al trabajo de titulación.

CLÁUSULA QUINTA. - EXCEPCIONES:

El Compromiso de Confidencialidad se excluye cuando:

- a) La información de propiedad institucional que sea de conocimiento y dominio público a efectos de cumplir con el principio de transparencia que dispone la Ley Orgánica de Transparencia y Acceso a la Información Pública; y,
- b) Por mandato judicial que exija su divulgación o disposición legal expresa.

CLÁUSULA SEXTA. - VIGENCIA:

El Compromiso de Confidencialidad de la información tendrá una vigencia indefinida, a partir de la fecha de su suscripción.

CLÁUSULA SÉPTIMA. - DECLARACIÓN:

El **Estudiante en proceso de titulación** declara que:

- a) La información de propiedad institucional y la de carácter personal de los pacientes del Hospital General del Norte de Guayaquil “Los Ceibos” y de los asegurados y empleadores del IESS constituye un activo intangible invaluable, por lo que, los riesgos por mal uso y/o divulgación indebida comporta que la entidad deba tomar medidas al respecto de la integridad de la misma;
- b) Conoce que todos los datos de la historia laboral o de la historia clínica que forman parte de la plataforma tecnológica institucional son de carácter reservado y confidencial respectivamente, por lo que utilizará la información exclusivamente de manera general a fin de resguardar la integridad de la misma;
- c) Mantendrá la información a la que tuviera acceso de manera confidencial y/o reservada, en virtud de la protección que gozan las personas por parte del orden jurídico constituido;



INSTITUTO ECUATORIANO DE SEGURIDAD SOCIAL

- d) El Hospital General del Norte de Guayaquil “Los Ceibos” y el IESS, cuando lo estimen pertinente y sin necesidad de notificación previa, podrán ejercer control y seguimiento de la información y/o datos que estén bajo su custodia, a fin de garantizar la confidencialidad, integridad y disponibilidad de la misma; y,
- e) Conoce las implicaciones que conlleva el incumplimiento de las cláusulas estipuladas en el presente documento y se compromete a su fiel cumplimiento. En caso de acción u omisión se someterá a las sanciones administrativas que le sean imputables, sin perjuicio de las responsabilidades civiles o penales a que hubiera lugar en virtud de la normativa legal vigente.

CLÁUSULA OCTAVA. - ACEPTACIÓN:

El Estudiante en proceso de titulación acepta el contenido de todas y cada una de las cláusulas del presente Compromiso de Confidencialidad de la Información, y en consecuencia se compromete a cumplirlas en toda su extensión, en virtud de lo cual y para los fines correspondientes, firma en dos ejemplares del mismo tenor y valor en la ciudad de Guayaquil a los 19 días del mes de enero del 2021.

Firma:

Cédula: 1103491435

Nombre: Roger Alexander Torres Yaguana

Anexo 11. Certificaciones del Hospital Ceibos.

Certificación

Dra. Priscilla Alcócer Cordero

**SUBDIRECTOR DE DOCENCIA DEL HOSPITAL GENERAL DEL NORTE DE
GUAYAQUIL “LOS CEIBOS”.**

Certifico:

Que el estudiante ***Roger Alexander Torres Yaguana***, con cédula de identidad N° 1103491435, de la Carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja, desarrolló en la Institución de Salud su trabajo de titulación denominado ***“APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS”***.

Guayaquil, 01 de marzo del 2021

A handwritten signature in blue ink is written over a faint, circular official stamp. The stamp contains the text 'Hospital General del Norte de Guayaquil' and 'Subdirector de Docencia'.

Dra. Priscilla Alcócer Cordero.

Subdirector de Docencia

Hospital General del Norte de Guayaquil “Los Ceibos

Certificación


Ing. Andrés Cobos M.

RESPONSABLE DEL ÁREA DE TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIONES DEL HOSPITAL GENERAL DEL NORTE DE GUAYAQUIL “LOS CEIBOS”.

Certifico:

Que el estudiante *Roger Alexander Torres Yaguana*, con cédula de identidad N° 1103491435, de la Carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja recopiló la información necesaria para su trabajo de titulación denominado ***“APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS”.***

Guayaquil, 01 de marzo del 2021

Instituto Ecuatoriano De Seguridad Social
HOSPITAL GENERAL DEL NORTE DE GUAYAQUIL LOS CEIBOS

Ing. Andrés Cobos M.
RESPONSABLE DEL ÁREA DE TECNOLOGÍAS DE LA
INFORMACIÓN Y COMUNICACIONES

Ing. Andrés Cobos M.

Director de TICs

Hospital General del Norte de Guayaquil “Los Ceibos”

Certificación

Dr. Antonio Javier Pozo García.

MÉDICO ESPECIALISTA DE MEDICINA INTERNA DEL HOSPITAL GENERAL DEL NORTE DE GUAYAQUIL “LOS CEIBOS”.

Certifico:

Que el estudiante *Roger Alexander Torres Yaguana*, con cédula de identidad N° 1103491435, probó la herramienta informática con mi persona, para su Proyecto de Titulación, denominado: ***“APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS”*** y bajo las correcciones pertinentes, se ejecutó y se está de acuerdo con la misma para su utilización.

Guayaquil, 17 de julio del 2021



Dr. Antonio Javier Pozo García
MÉDICO ESPECIALISTA DE MEDICINA INTERNA
C.I. 1096-201E-1752578 C.I. 0920915750
#E-5 Instituto Ecuatoriano De Seguridad Social
HOSPITAL GENERAL DEL NORTE DE GUAY. LOS CEIBOS

Dr. Antonio J. Pozo G.

MÉDICO ESPECIALISTA

Certificación

Dr. Lenin Francisco Torres Yaguana.

MÉDICO ESPECIALISTA DE MEDICINA INTERNA DEL HOSPITAL GENERAL DEL NORTE DE GUAYAQUIL “LOS CEIBOS”.

Certifico:

Que el estudiante *Roger Alexander Torres Yaguana*, con cédula de identidad N° 1103491435, probó la herramienta informática con mi persona, para su Proyecto de Titulación, denominado: **“APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS”** y bajo las correcciones pertinentes, se ejecutó y se está de acuerdo con la misma para su utilización.

Guayaquil, 17 de julio del 2021



Dr. Lenin Francisco Torres Yaguana-
MEDICINA INTERNA
1006-2016-1752577 C.I. 1103504690
+ESS Instituto Ecuatoriano De Seguridad Social
HOSPITAL GENERAL DEL NORTE DE GUAY. LOS CEIBOS

Dr. Lenin F. Torres Y.

MÉDICO ESPECIALISTA

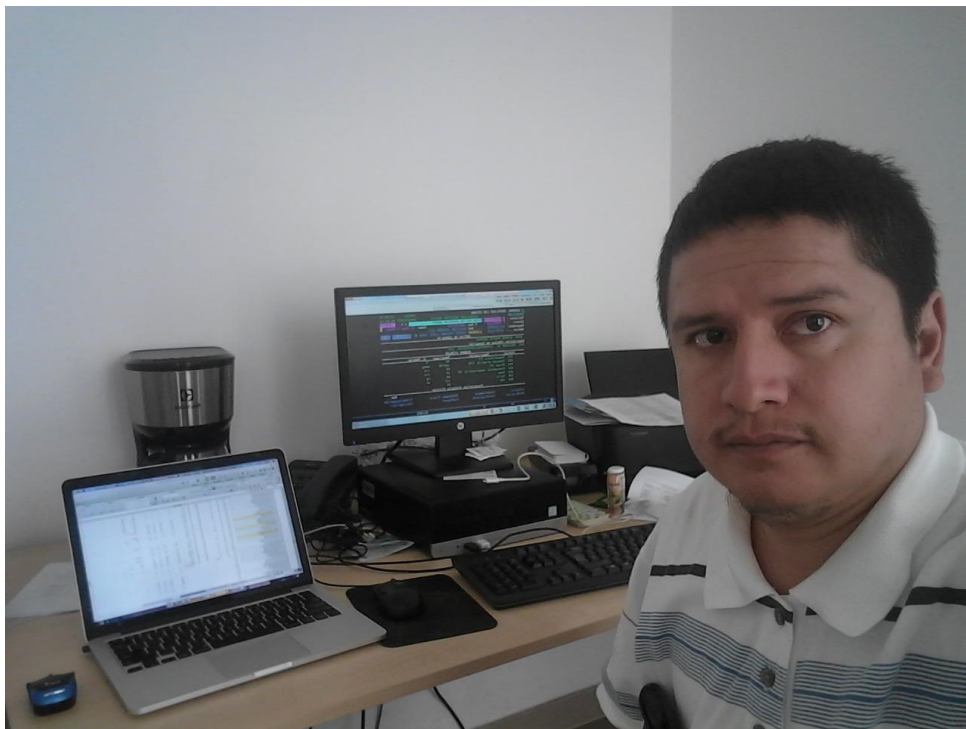
Anexo 12: Fotografías de la aprobación de la Aplicación Web.



Dr. Antonio Pozo – Roger Torres



Dr. Lenin Torres – Roger Torres



Recopilación manual de la información de los pacientes – Roger Torres

Anexo 13: Manuales del programador y usuario.

Manual del Programador

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos

Introducción

El presente manual refiere los aspectos informáticos del sistema web CardioZT. El manual familiariza al personal técnico especializado delegado de las actividades de mantenimiento, revisión, solución de problemas, instalación y configuración del sistema.

Información destacada

El manual técnico hace referencia a la información necesaria con el fin de orientar al personal en la concepción, planteamiento, análisis, programación e instalación del sistema. Es de notar que la redacción propia del manual técnico está orientada a personal con conocimientos en sistemas y tecnologías de información, conocimientos de programación avanzada sobre entorno web, administración de bases de datos, responsables del mantenimiento e instalación del sistema en los servidores.

Objetivos

Instruir el uso adecuado del Sistema Web CardioZT, para el acceso oportuno y adecuado en la instalación del mismo, mostrando los pasos a seguir en el proceso de instalación, así como la descripción de los archivos relevantes del sistema los cuales nos orienten en la configuración y soporte del mismo.

1. Requerimientos

El sistema puede ser instalado en cualquier sistema operativo que cumpla los siguientes requerimientos:

- Framework Web Django 1.11 o superior
- Python 3.5.5 o superior
- SQLite3 o superior
- Navegador Web

2. Instalación y Configuración

De los diferentes métodos de instalación existentes, a continuación, se narra una forma sencilla y segura de realizar la instalación, la cual consiste en la siguiente manera de realizarla.

Esta es la opción más adecuada a causa de que es la más oportuna, sin tener que tocar código alguno. Al mismo tiempo, proporciona seguridad y es la que se sugiere debido a su practicidad en el framework django.

El método a continuación está enfocado a Sistema Linux, en el cual se tiene como requisito previo Python3.5.5 o Python3.5.9.

1. Instalar gestor de paquetes si no existe:

```
sudo apt install python3-pip
```

2. Instalar virtualenv para crear un entorno virtual de la app:

```
pip3 install virtualenv --user
```

3. Crear carpeta para almacenar el entorno virtual:

```
mkdir ~/virtualenv
```

4. Crear el entorno virtual:

```
python3 -m virtualenv ~/virtualenv/cardioEnv
```

5. Activar el entorno virtual:

```
source ~/virtualenv/cardioEnv/bin/activate
```

6. Entrar en la carpeta del proyecto e instalar los requerimientos de la app:

```
pip install -r requirements.txt
```

A continuación, para inicializar la aplicación CardioZT, se lo realiza con los siguientes pasos.

1. Para iniciar el servicio se tiene que entrar a la carpeta `proyectoCardio` y ejecutar:

```
python manage.py runserver
```

2. Entrar en admin y agregar “/admin” a la url (Ejemplo: 127.0.0.1/admin):

```
usuario: admin
```

```
contraseña: Pass1234
```

3. Diagrama de Clases

A continuación, se muestra el diagrama de clases de las tablas que componen el sistema:

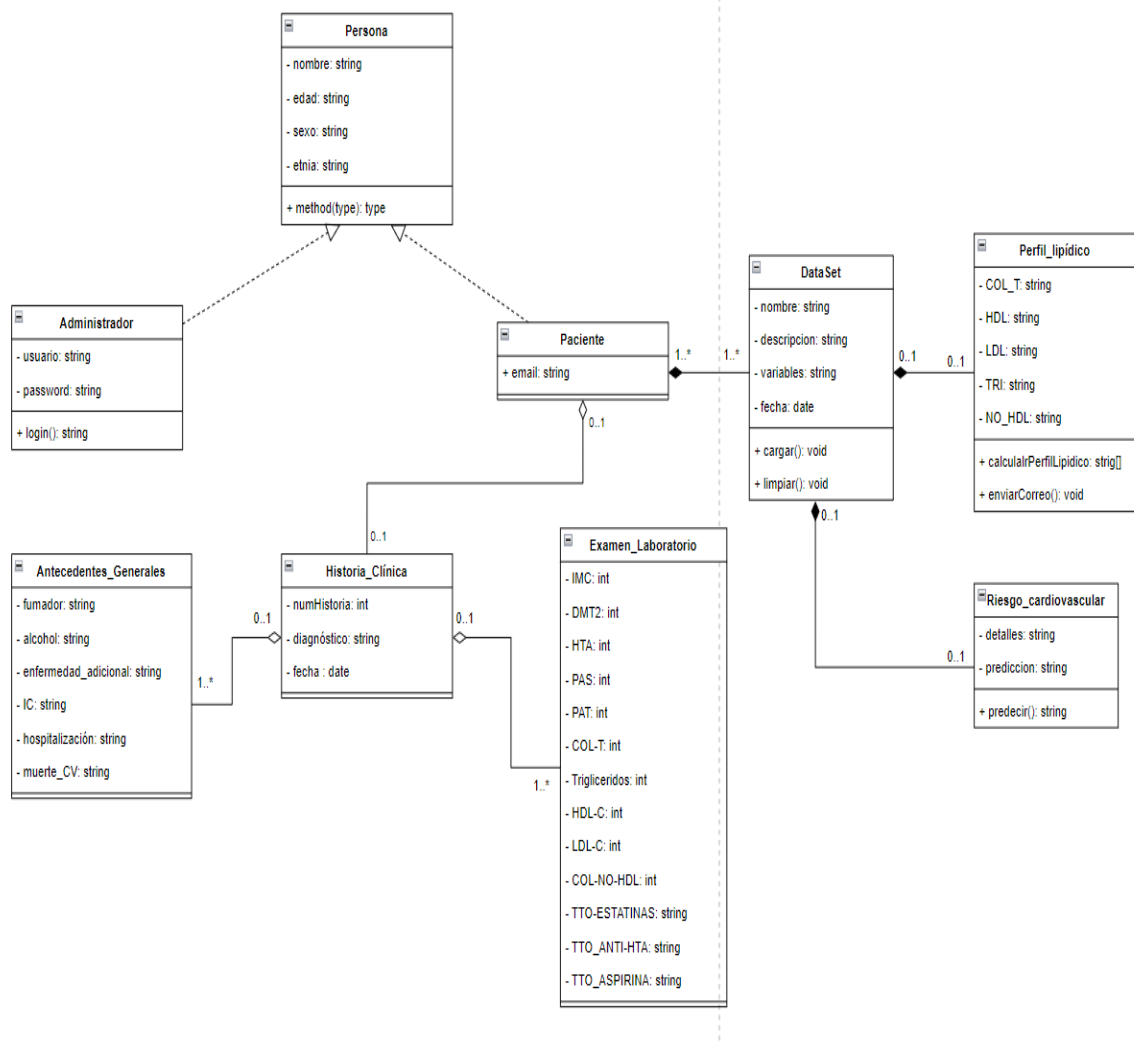


Fig. 1 Diagrama de Clases

Configuración del sistema

El sistema CardioZT está desarrollado bajo el framework Django en su versión 1.11, el cual da la facilidad de estructurar el sistema de manera que se facilita el mantenimiento a dicha solución, a continuación, se describe la estructura básica del sistema y se enfatiza en los archivos y directorios relevantes para su configuración y adaptación.

Estructura Raíz

El proyecto CardioZT tiene la siguiente estructura de directorios:

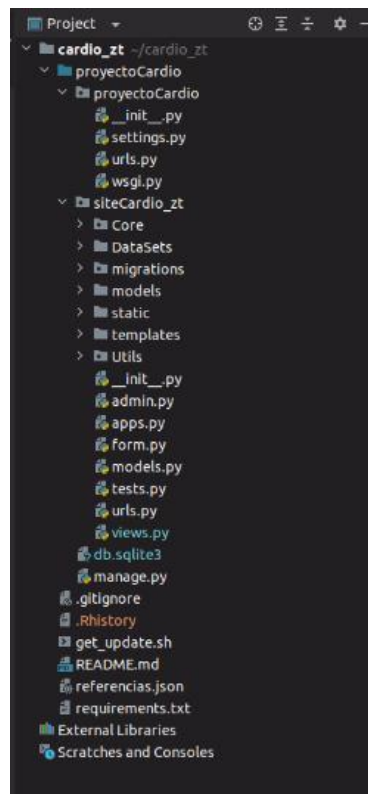


Fig. 2 Estructura raíz del proyecto CardioZT

A continuación, describiremos los directorios y archivos más importantes:

Directorio siteCardio_zt

siteCardio_zt es usado para ofrecer un hogar por defecto a todo el código principal del proyecto. Eso incluye clases que puedan ofrecer funcionalidad a la aplicación y más. Es considerado el directorio más importante del proyecto.

El directorio siteCardio_zt tiene a su vez otros subdirectorios importantes, pero uno de los más utilizados es el directorio Core en el cual ubicaremos las clases más importantes

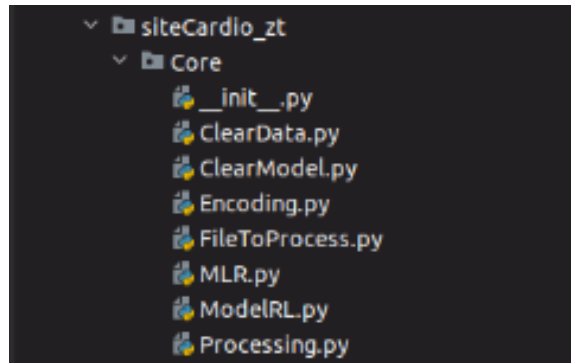


Fig. 3 Directorio siteCardio_zt

- ClearData: Contiene las dos clases principales para la creación de objetos y limpieza de los datos automática de los archivos estáticos csv.
- FileToProcess: clase que representa un archivo csv para ser procesado como un objeto
- MLR: clase del modelo de aprendizaje automático con regresión lineal múltiple.

Directorio DataSets

En el subdirectorio ClearData, el sistema contiene los archivos estáticos csv para proceder a ejecutar la aplicación con los archivos ya limpios.

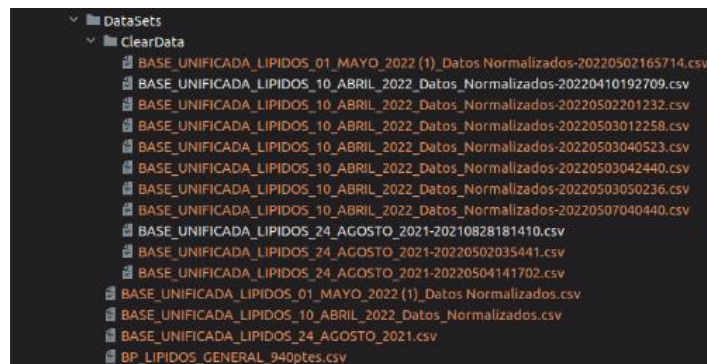


Fig. 4 Directorio DataSets

El archivo `BASE_UNIFICADA_LIPIDOS_24_AGOSTO_2021-20210828181410.csv` es uno de los archivos ya limpiado automáticamente para su ingreso al modelo predictivo con regresión lineal múltiple y este a su vez hace la relación entre la plantilla y la vista.

Directorio templates

Dentro de este directorio se encuentran las pantallas de la aplicación, en el cual las más importantes son:

- **create_model.html:** En este archivo se acciona el modelo predictivo para el pronóstico del riesgo cardiovascular.
- **list_data.html:** En este archivo se encuentran los datasets para el funcionamiento correcto del modelo predictivo.
- **limpiar_datos.html:** En este archivo se realiza la limpieza automática de los datos.
- **perfil_lipidico.html:** En este archivo se hace el cálculo matemático del perfil lipídico del paciente.

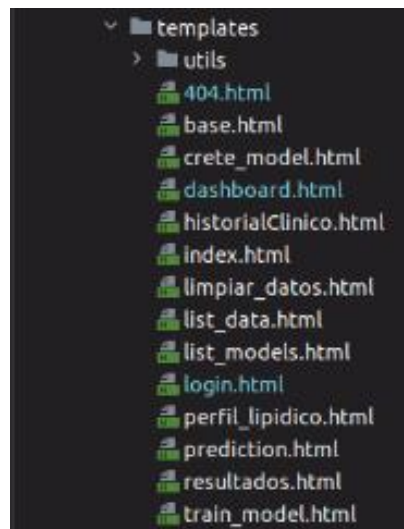


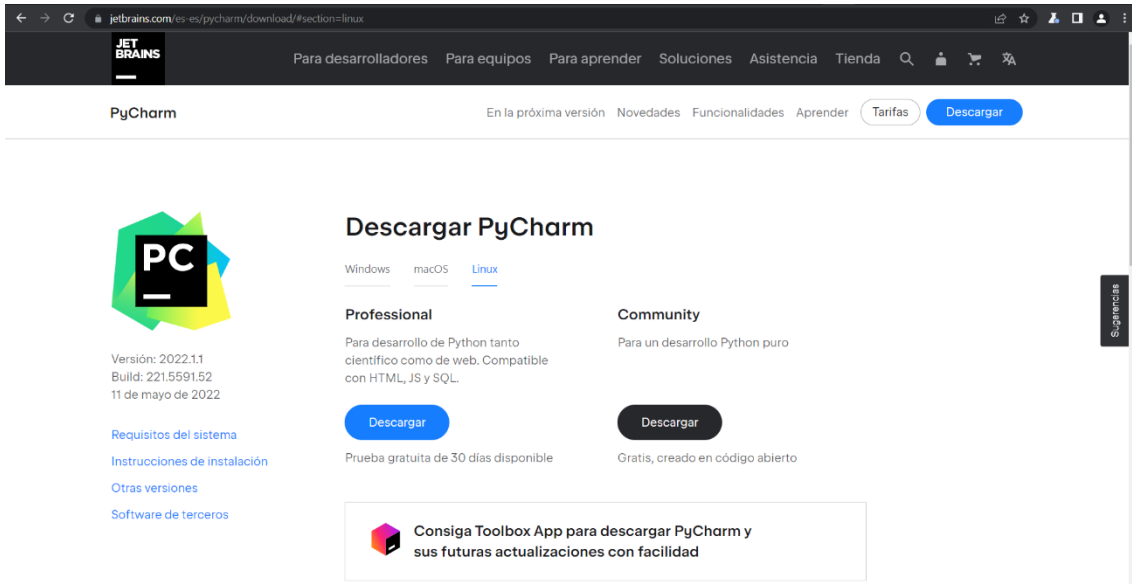
Fig. 5 Directorio Templates

Todos los archivos y directorios no mencionados son parte importante para el funcionamiento del sistema, no se hacen referencia en este documento debido a que solo se enfatizan los archivos que el usuario puede en un caso dado modificar, con conocimiento previo de lo que se hace.

4. Desarrollo Web

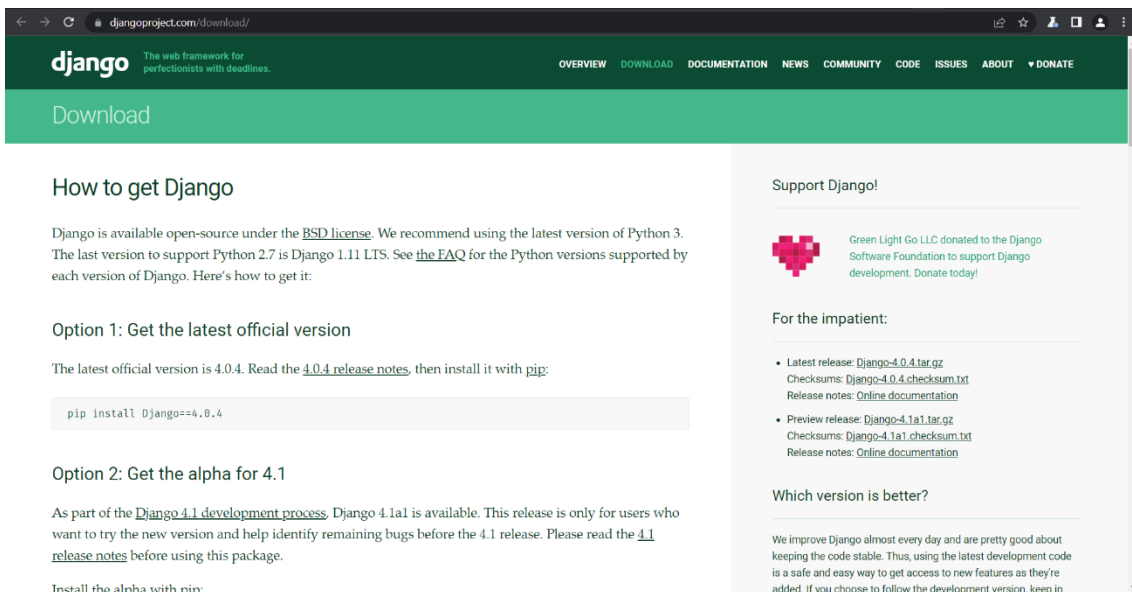
Para el desarrollo del aplicativo web se utilizó el IDE Pycharm Community, en el cual fueron necesarios configurar varios elementos. Se trabajó en el lenguaje de programación Python en el sistema operativo Linux, estas distribuciones están disponibles en:

<https://www.jetbrains.com/es-es/pycharm/download/#section=linux>



The screenshot shows the JetBrains website's download page for PyCharm on Linux. The page features the PyCharm logo (a green and yellow 'PC' icon) and provides version information: 'Versión: 2022.1.1', 'Build: 221.5591.52', and '11 de mayo de 2022'. It lists system requirements, installation instructions, and other versions. Two download options are presented: 'Professional' (for Python development, compatible with HTML, JS, and SQL) and 'Community' (for pure Python development). Both options have 'Descargar' buttons. A promotional banner at the bottom encourages downloading the Toolbox App for easier updates.

Después de instalado el IDE pertinente al sistema operativo, es necesario realizar la descarga y configuración del Framework Django, para esto se debe dirigir a la URL: <https://www.djangoproject.com/download/> donde se encontrará la opción: Descargar Django



The screenshot shows the Django project's download page. It features the Django logo and a navigation menu with links for Overview, Download, Documentation, News, Community, Code, Issues, About, and Donate. The main heading is 'Download'. The page provides instructions on how to get Django, including the license (BSD) and the recommended Python version (3). It offers two options: 'Option 1: Get the latest official version' (4.0.4) and 'Option 2: Get the alpha for 4.1'. A code block shows the command: `pip install Django==4.0.4`. The page also includes a 'Support Django!' section with a donation link and a 'For the impatient:' section with links to the latest and preview releases, including their checksums and release notes. A 'Which version is better?' section explains the development process and encourages users to keep the code stable.

Después de realizada la instalación procedemos a crear un nuevo proyecto. El nombre del proyecto es proyectoCardio y la aplicación será siteCardio_zt, el cual contendrá a la carpeta Core.

Luego en la carpeta Core se crea Clear.Data.py, la cual se compone de las clases ClearColumns y ClearData. En la clase ClearColumns se crea objetos de una columna de un csv para proceder a limpiar, en la cual se tiene una función para validar el tipo de opción que se va a aplicar para limpiar la columna (por defecto es: delete). Los tipos soportados son:

- replace: reemplaza los valores vacíos con otro valor (se tiene que especificar: replace_with).
- auto_complete: auto completa los valores vacíos tomando en cuenta los valores existentes.
- delete: elimina los valores vacíos.

```
class ClearColumns:
```

```
    clear_options = ["replace", "auto_complete", "delete"]
```

```
    def __init__(self, column_name, clear_type="delete", replace_with=None):
```

```
        self.column_name = column_name
```

```
        self.clear_type = self.option_to_clear(clear_type)
```

```
        self.replace_with = replace_with
```

```
    def option_to_clear(self, option_type) -> str:
```

```
        if option_type is None:
```

```
            return self.clear_options[2]
```

```
        elif option_type in self.clear_options:
```

```
            return option_type
```

```
        else:
```

```
            raise Exception("Option not support! Only support: {}".format(self.clear_options))
```

En la clase ClearData, se limpian los datos de un csv, en la cual se tiene una función para validar que todos los objetos de la lista son de tipo: ClearColumns. Además, una función para reemplazar los valores vacíos de una columna, otra función para eliminar los valores vacíos de una columna, otra para limpiar las cabeceras automáticamente, otra para enumerar las cabeceras automáticamente, otra para establecer el valor de la variable: index_to_clear, otra para limpiar las cabeceras del archivo csv, otra función para empezar con la limpieza del archivo y una función para guardar los cambios realizados en el archivo.

```
class ClearData:

    actions_clear_headers = ["auto", "enumeration"]

    def __init__(self, file_process: FileToProcess, save_path: str = None, index_to_clear: list = None):

        self.file_process = file_process

        self.save_path = save_path

        self.index_to_clear = self.validate_list_index(index_to_clear) if index_to_clear else None

        self.df_file_process = file_process.get_pandas_file()

    @staticmethod

    def validate_list_index(list_clear_columns):

        if all(isinstance(obj, ClearColumns) for obj in list_clear_columns):

            return list_clear_columns

        else:

            raise Exception("Type list not support!")

    def __replace(self, column: ClearColumns):

        self.df_file_process[column.column_name].fillna(column.replace_with, inplace=True)

    def __auto_complete(self, column: ClearColumns):

        pass

    def __delete(self, column: ClearColumns = None):

        if column:

            print("Delete: {}".format(column.column_name))
```

```

self.df_file_process = self.df_file_process.dropna(subset=[column.column_name])

else:

    print("Delete all")

    self.df_file_process = self.df_file_process.dropna()

def __auto_clear_headers(self, current_headers):

    rename_headers = []

    for index, header in enumerate(current_headers):

        if "Unnamed:" in header:

            rename_headers.append("column_" + str(index))

        else:

            rename_headers.append(str(header).strip().replace(" ", "_"))

    columns_to_rename = dict(zip(current_headers, rename_headers))

    logger.info("Data:")

    logger.log(level_log, "data rename: {}".format(rename_headers))

    logger.log(level_log, "")

    logger.log(level_log, "data current: {}".format(current_headers))

    logger.log(level_log, "")

    logger.log(level_log, "data zip : {}".format(columns_to_rename))

    logger.info("Data:")

    self.df_file_process = self.df_file_process.rename(columns=columns_to_rename)

def __enumeration_clear_headers(self, current_headers):

    for index, header in enumerate(current_headers):

        current_headers[index] = "column_" + str(index)

    self.df_file_process = self.df_file_process.rename(columns=current_headers)

def set_index_to_clear(self, values: list):

    if values:

```

```

        self.index_to_clear = values

    else:

        raise Exception("Please, send values!")

def clear_headers(self, custom_headers=None, action="auto"):

    current_headers = self.file_process.get_headers()

    if custom_headers is None:

        if action == "auto":

            self.__auto_clear_headers(current_headers)

        elif action == "enumeration":

            self.__enumeration_clear_headers(current_headers)

        else:

            raise Exception("Invalid action!")

    else:

        self.df_file_process.rename(columns=custom_headers)

def clear_columns(self):

    if self.index_to_clear:

        for column in self.index_to_clear:

            if column.clear_type == "replace":

                self.__replace(column)

            elif column.clear_type == "auto_complete":

                self.__auto_complete(column)

            elif column.clear_type == "delete":

                self.__delete(column)

        else:

            raise Exception("Error: index_to_clear is NONE.")

def delete_nan(self):

    self.__delete()

```

```

def write_changes(self):

    if self.save_path:

        path_to_save = self.save_path

    else:

        print("NO PATH!!!")

        path_to_save = dirname(self.file_process.path_file)

    self.file_process.file_name = self.file_process.file_name + "-" + strftime("%Y%m%d%H%M%S",
gmtime())

    self.file_process.path_file = os.path.join(path_to_save, self.file_process.file_name +
self.file_process.extension_name)

    print("Path to save: {}".format(self.file_process.path_file))

    return self.file_process.path_file if self.df_file_process.to_csv(self.file_process.path_file) else None

def encoding_data(self, data_to_clear: [Replace]):

    replace = dict()

    for data in data_to_clear:

        replace.update(data.get_data_replace())

    print("Replace data: {}".format(replace))

    self.df_file_process.replace(replace, inplace=True)

def cardio_risk(self, **kwargs):

    l = calc_l(gender="H" if get_gender(kwargs.get("sexo")) else "M", edad=kwargs.get("edad"),
colesterol=kwargs.get("colesterol"),

        HDL_C=kwargs.get("hdl_col"), PAS=kwargs.get("pas"), PAD=kwargs.get("pad"),

        diabetes=kwargs.get("diabetes"), fumador=kwargs.get("fumador"))

    g = get_g(gender="H" if get_gender(kwargs.get("sexo")) else "M")

    b = calc_b(l, g)

    s = get_s(gender="H" if get_gender(kwargs.get("sexo")) else "M")

    return calc_r(s, b)

def calc_cardio_risk(self):

```



```

self.df_file_process["CARDIO_RISK"] = self.df_file_process.apply(lambda x: self.cardio_risk(
    sexo=x["SEXO"], edad=x["EDAD"], colesterol=x["COL-T"], hdl_col=x["HDL-C"],
    pas=x["PAS"], pad=x["PAD"],
    diabetes=x["DMT2"], fumador=["FUMADOR"]), axis=1)

print("Function apply")

return True

def calc_perfil(self):

    list_to_calc = ["COLUM-NAME", "COLUM-NAME"]

    for a in list_to_calc:

        print("Function apply")

    return True

def overwrite(self):

    pass

```

En la clase FileToProcess, se crea una clase que representa un archivo csv para ser procesado como un objeto, en la cual se obtiene una lista de las cabeceras (la primera columna) del archivo que retorna una lista de string. Tiene una función para obtener el número de columnas del archivo csv. Además, una función para obtener el número de filas del archivo csv, una función para tener el archivo csv como un objeto de tipo pandas y una función para convertir el csv en un objeto de tipo numpy.

```

class FileToProcess:

    def __init__(self, path_file, delimiter=",", index_column=0):

        self.path_file = path_file

        self.delimiter = delimiter

        self.__pandas_file = pd.read_csv(self.path_file, delimiter=self.delimiter, index_col=index_column)

        self.file_name = get_basename(path_file)["file_name"]

        self.extension_name = get_basename(path_file)["file_extension"]

        self.headers = self.get_headers()

```

```

self.columns_number = self.get_columns_number()

self.row_number = self.get_row_numbers()

def get_headers(self) -> list:

    if self.extension_name == ".csv":

        return list(self.__pandas_file.columns)

    else:

        raise Exception("Not supported extension.")

def get_columns_number(self) -> int:

    return self.__pandas_file.shape[1]

def get_row_numbers(self) -> int:

    return self.__pandas_file.shape[0]

def get_pandas_file(self) -> pd:

    return self.__pandas_file

def convert_to_numpy(self):

    return self.__pandas_file.to_numpy(copy=True)

def get_columns_numpy(self, columns):

    return self.__pandas_file[columns].to_numpy(copy=True)

```

En la clase MLR, la cual es la clase de la regresión lineal múltiple, se definen los datos correspondientes a las etiquetas, se separa los datos en entrenamiento y prueba para probar el algoritmo, se define el algoritmo a utilizar, se entrena el modelo y se realiza la predicción.

```

class MLR:

    def __init__(self,

        file_to_process: FileToProcess = None,

        column_name_x: list = None,

        column_name_y: list = None,

```



```

print(type(self.model_rl))

# Entreno el modelo

self.model_rl.fit(X_train, y_train)

# Realizo una predicción

Y_pred_multiple = self.model_rl.predict(X_test)

print('PREDICCIÓN DEL MODELO DE REGRESIÓN LINEAL MULTIPLE')

print(Y_pred_multiple)

print('DATOS DEL MODELO REGRESIÓN LINEAL MULTIPLE')

print()

print("Valor de las pendientes o coeficientes "a":")

print(self.model_rl.coef_)

print("Valor de la intersección o coeficiente "b":")

print(self.model_rl.intercept_)

print('Precisión del modelo:')

self.accuracy = self.model_rl.score(X_train, y_train)

print(self.model_rl.score(X_train, y_train))

def prediction(self, x_test, model: RegressionLinearModel):

    lr = self._get_model(model)

    p = lr.predict(x_test)

    pr = PredictionsResult(coef=lr.coef_, id_regresion_lm=model, intercept=lr.intercept_, prediction=p)

    pr.save()

    return pr

def train_model(self, x_multiple, y_multiple, model=None):

    # Separo los datos de "train" en entrenamiento y prueba para probar el algoritmo

    X_train, X_test, y_train, y_test = train_test_split(x_multiple, y_multiple, test_size=self.test_size)

    if model:

        model.fit(X_train, y_train)

```

```

        self.model_rl = model

    else:

        self.model_rl.fit(X_train, y_train)

    print('Precisión del modelo:')

    self.accuracy = self.model_rl.score(X_train, y_train)

    print(self.model_rl.score(X_train, y_train))

    print("OK")

def create_model(self):

    x_multiple = self.file_to_process.get_columns_numpy(self.column_name_x)

    y_multiple = self.file_to_process.get_columns_numpy(self.column_name_y)

    self.model_rl = linear_model.LinearRegression()

    self.train_model(x_multiple, y_multiple)

def _write_model(self, extension=".sav"):

    self.file_name = "model" + "-" + strftime("%Y%m%d%H%M%S", gmtime()) + extension

    file = os.path.join(settings.BASE_DIR, self.model_path, self.file_name)

    joblib.dump(self.model_rl, file)

def save_model(self, file, name, description):

    if self.model_rl:

        self._write_model()

        model_rl = RegressionLinearModel()

        model_rl.id_file = file

        model_rl.name = name

        model_rl.file_name = self.file_name

        model_rl.model_path = self.model_path

        model_rl.dependent_var = self.column_name_x

        model_rl.independent_var = self.column_name_y

```

```
model_rl.description = description

model_rl.coef = list(self.model_rl.coef_)

model_rl.accuracy = self.accuracy

model_rl.intercept = self.model_rl.intercept_.tolist()

model_rl.save()

else:

    raise Exception("The model does not exist (pls, train the model first)")
```

Manual del Usuario

Proyecto: Aplicación web determinar el perfil lipídico y su relación con el riesgo cardiovascular en pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos

Aplicación web CardioZT

El sistema web CardioZT, es una aplicación desarrollada con el propósito de facilitar la información necesaria del estado de salud de los pacientes diabéticos tipo 2 e hipertensos de la Consulta Externa del Hospital Ceibos de Guayaquil. El presente documento permitirá ingresar las variables necesarias para hacer la predicción del riesgo cardiovascular y el cálculo matemático del conjunto de variables para determinar los niveles del perfil lipídico del paciente.

Interfaces del Sistema

En lo referente a las interfaces del sistema web CardioZT, son las siguientes: el inicio de sesión, el listado de los datos, la limpieza de los datos, el modelo predictivo para el riesgo cardiovascular y el cálculo del perfil lipídico del paciente con diabetes mellitus tipo 2 e hipertensión arterial del hospital.

Primeramente, el ingreso a la aplicación web CardioZT, se puede realizar en cualquiera de los 3 navegadores preferidos: Firefox, Chrome, Google Edge. Utilizando la url del proyecto: <https://hmswp8zzmf.us-east-1.awsapprunner.com/>

Seguidamente al cargar el sistema con la url de la aplicación, se muestra la pantalla de inicio de sesión, en la cual se deberá ingresar un usuario y una contraseña del administrador

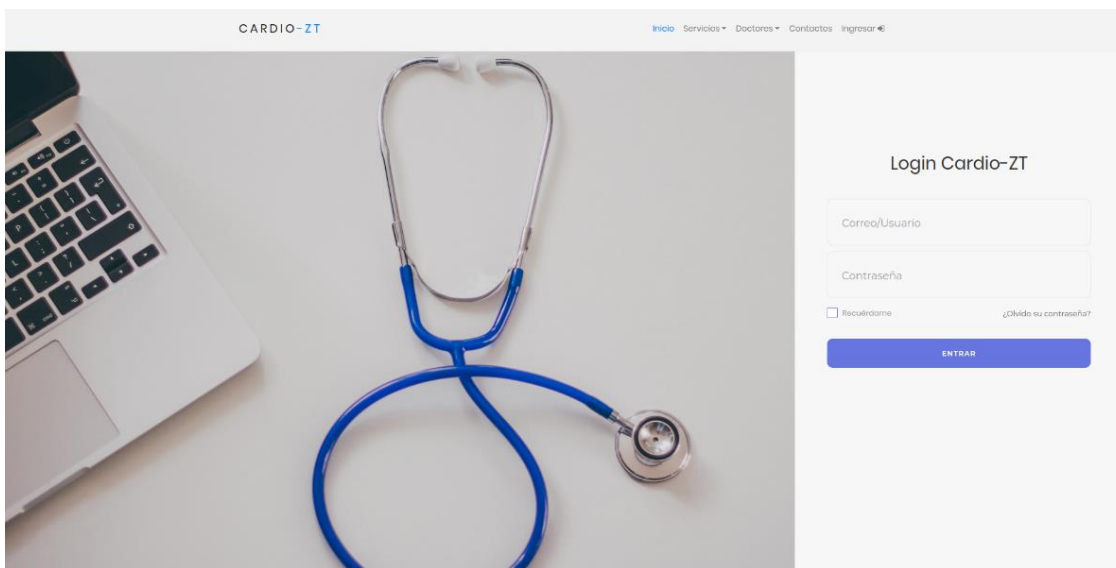


Fig. 6 Inicio de sesión del CardioZT

Luego una vez logueado con el usuario administrador se ingresa a la aplicación, en la cual se muestran los 4 módulos del sistema: modelos, listar datos, limpiar datos y perfil lipídico.

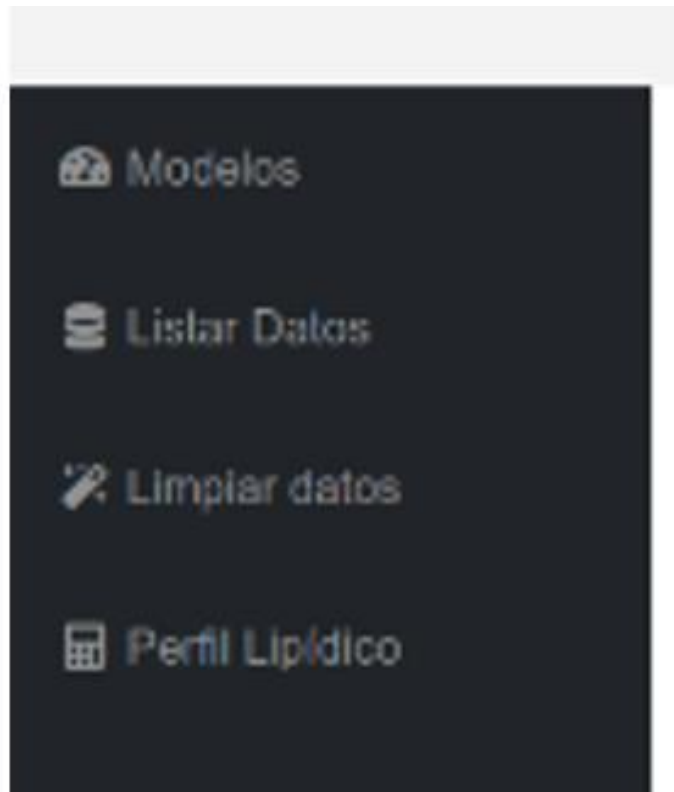


Fig. 7 Items del proyecto CardioZT

Modelos: Contiene el conjunto de variables para el modelo predictivo que el usuario administrador va a seleccionar para poder obtener el riesgo cardiovascular del paciente.

Listar Datos: Posee todos los datasets cargados de la aplicación con su respectivo nombre, descripción, y fecha de creación.

Limpiar datos: Incluye la limpieza automática del dataset seleccionado para poder dejarlo listo para el ingreso correcto al modelo predictivo como también al cálculo del perfil lipídico del paciente.

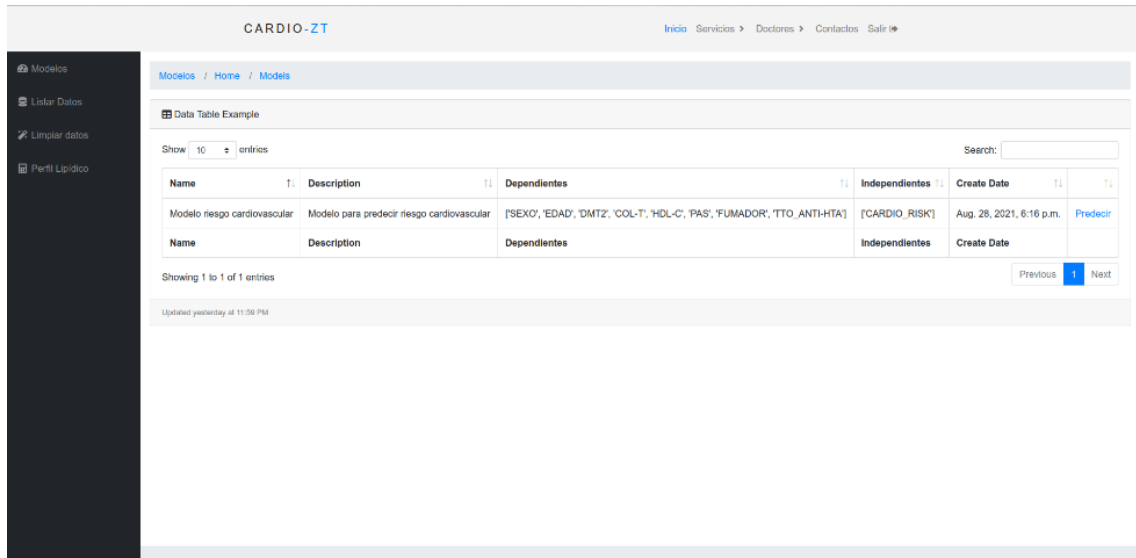
Perfil lipídico: Abarca los valores de las variables utilizadas para el cálculo matemático de los niveles del perfil lipídico del paciente.

Pantalla principal del proyecto CardioZT

Esta es la pantalla principal del sistema, la cual está conformada por los módulos de la aplicación y el modelo predictivo para el pronóstico del riesgo cardiovascular, cuyos

campos son nombre, descripción, las variables independientes y dependiente y la fecha de creación.

Se acciona el botón predecir y se selecciona las variables independientes e independiente para el riesgo cardiovascular con su porcentaje de entrenamiento, para hacer la predicción.



The screenshot shows the main interface of the Cardio-ZT application. At the top, there is a navigation bar with the title 'CARDIO-ZT' and links for 'Inicio', 'Servicios', 'Doctores', 'Contactos', and 'Salir'. A sidebar on the left contains menu items: 'Modelos', 'Listar Datos', 'Limpiar datos', and 'Perfil Lipídico'. The main content area displays a 'Data Table Example' with a search bar and a table. The table has columns for Name, Description, Dependientes, Independientes, and Create Date. A single entry is shown for 'Modelo riesgo cardiovascular' with a 'Predecir' button. Below the table, it indicates 'Showing 1 to 1 of 1 entries' and 'Updated yesterday at 11:59 PM'.

Name	Description	Dependientes	Independientes	Create Date
Modelo riesgo cardiovascular	Modelo para predecir riesgo cardiovascular	[SEXO, 'EDAD', 'DMTZ', 'COL-T', 'HDL-C', 'PAS', 'FUMADOR', 'TTO ANTI-HTA]	[CARDIO_RISK]	Aug. 28, 2021, 6:16 p.m.

Fig. 8 Pantalla principal del CardioZT

Pantalla de administración del riesgo cardiovascular

En esta pantalla se observa la estructura del modelo predictivo para el riesgo cardiovascular, el cual consta de las variables: sexo, la edad, diabetes, el colesterol total, la lipoproteína de alta densidad, la presión arterial sistólica, fumador y tratamiento antihipertensivo como variables predictoras para el modelo.

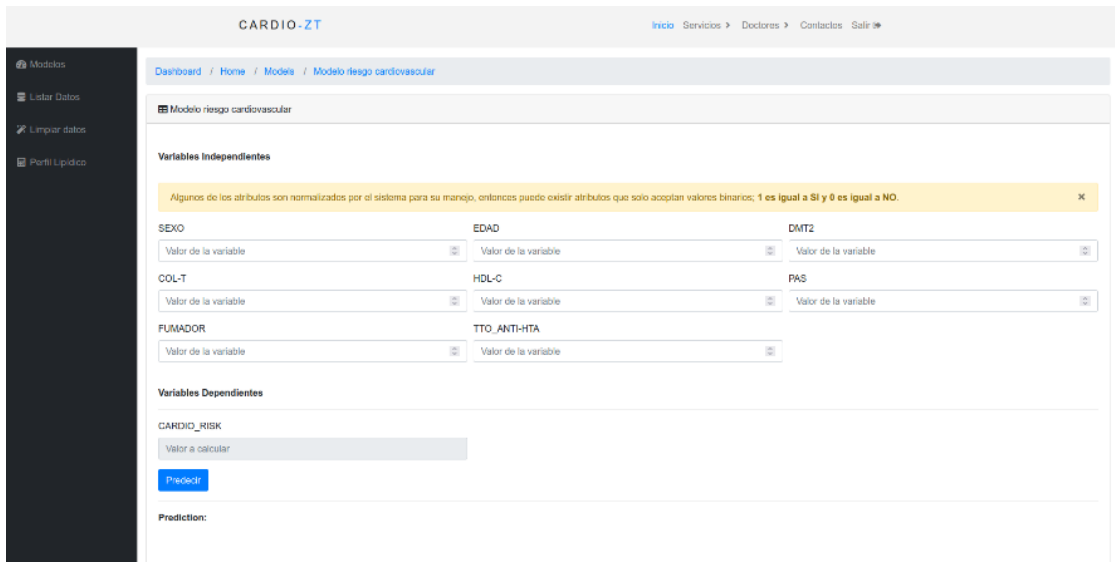


Fig. 9 Pantalla de administración del riesgo cardiovascular

Pantalla de obtención del riesgo cardiovascular

En esta pantalla se observa el resultado del riesgo cardiovascular del paciente analizado con el modelo de la regresión lineal múltiple, en la cual se obtiene el porcentaje del riesgo y la etiqueta del estado que se encuentra el paciente: bajo, moderado, alto y muy alto riesgo.

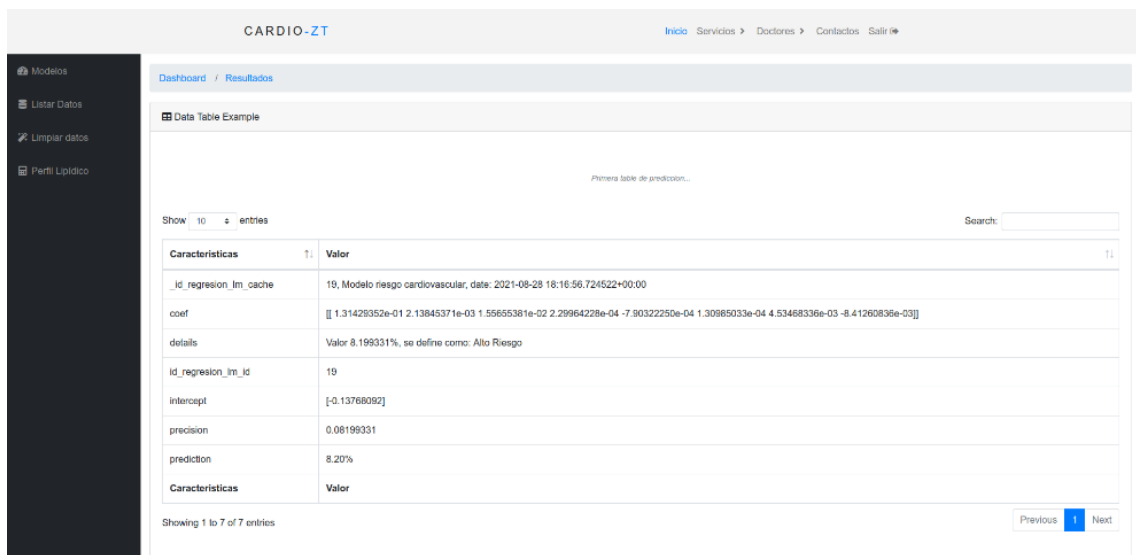


Fig. 10 Pantalla de obtención del riesgo cardiovascular

Pantalla de creación de modelo automático

En esta pantalla el usuario podrá seleccionar las variables para la creación del modelo predictivo del riesgo cardiovascular, que constará de un nombre, un porcentaje de entrenamiento, una descripción, las variables independientes y la variable dependiente para realizar la predicción con aprendizaje automático supervisado.

The screenshot displays the 'New Model' creation interface in the 'CARDIO-ZT' application. The page title is 'CARDIO-ZT' and the breadcrumb is 'Inicio Servicios > Doctores > Contactos Salir'. The main content area is titled 'BASE_UNIFICADA_LIPIDOS_24_AGOSTO_2021-20210828181410'. It contains several form fields: 'Nombre del modelo' with a text input 'Nombre' and a note 'No mayor a 60 caracteres.'; 'Porcentaje de entrenamiento' with a slider set to 20% and a note 'No debe ser superior al 50%.'; 'Descripción' with a large text area; 'Variables dependientes' with a text input 'VARIABLE, VARIABLE2' and a note 'Los nombres deben ser separados por comas.'; and 'Variables independientes' with a list of checkboxes for variables like NOMBRE_PACIENTE, SEXO, EDAD, DMT2, HTA, COL-T, TRIGLICERIDOS, HDL-C, LDL-C, COL_NO-HDL, RAZA, PAS, PAD, IMC, FUMADOR, ALCOHOL, TTO_ESTATINAS, TTO_ANTI-HTA, TTO_ASPIRINA, ENFERMEDAD_ADICIONAL, IC, HOSPITALIZACION, MUERTE_CV, and CARDIO_RISK. A 'Crear Modelo' button is at the bottom left, and the footer says 'Created Aug. 28, 2021'.

Fig. 11 Pantalla de creación del modelo automático

Pantalla principal del listado de datos

En esta pantalla se visualiza los datasets cargados en el sistema con los que realiza el aplicativo su funcionamiento, para el modelo predictivo del riesgo cardiovascular y para el cálculo del perfil lipídico, la cual contiene un nombre, una descripción y las columnas con las que están compuestos los datos.

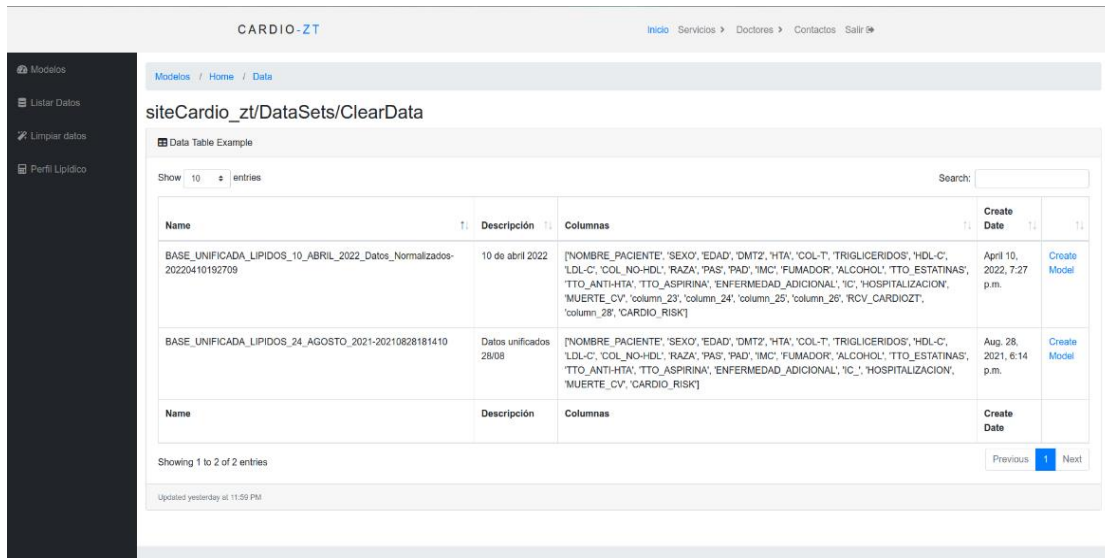


Fig. 12 Pantalla principal del listado de datos

Pantalla de administración de limpieza de datos

En esta pantalla el usuario puede seleccionar el dataset a pasar la limpieza de datos de forma automática donde se selecciona el dataset, que consta de un nuevo nombre, una descripción, el tipo de limpieza tanto para el riesgo cardiovascular como para el perfil lipídico y un botón de limpiar que devolverá un objeto Jason, en el cual se visualizará las columnas limpias de forma automática y el total de registros del dataset.

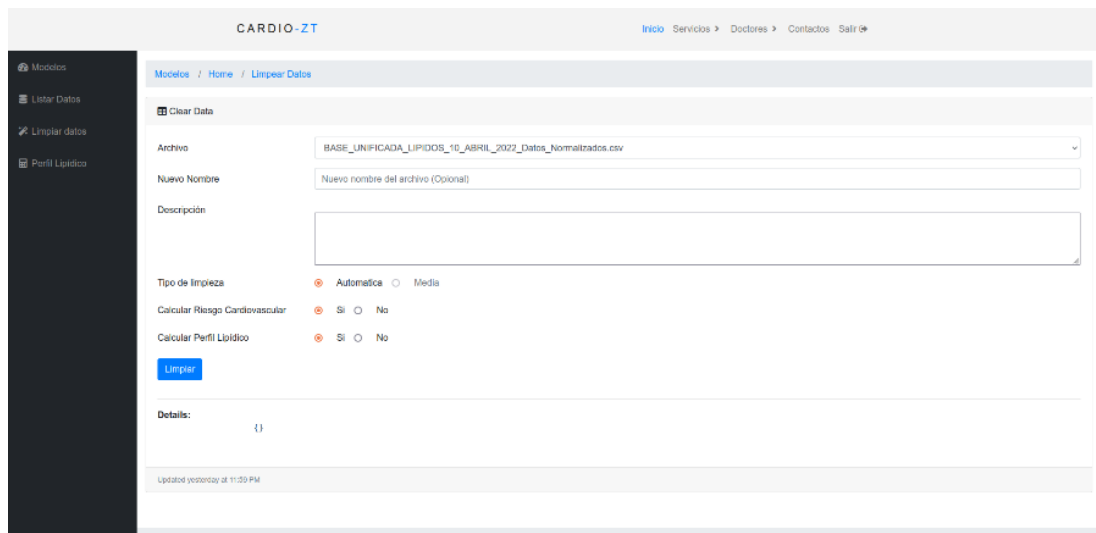


Fig. 13 Pantalla de administración de limpieza de datos

Pantalla de Estadísticas de los Pacientes

En esta pantalla se visualizan las estadísticas del conjunto total de registros de pacientes del dataset analizado, en la cual se observa el porcentaje de géneros masculino y femenino, muerte y rangos de edad.

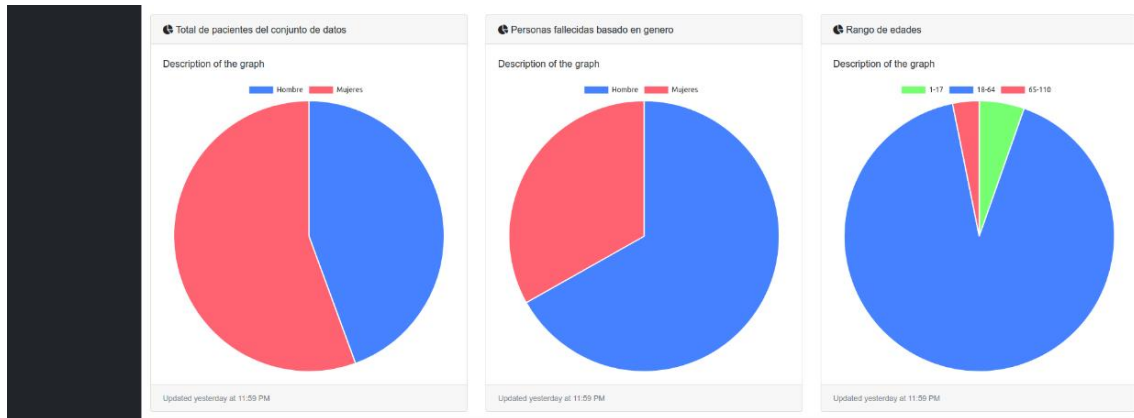


Fig. 14 Pantalla de Estadísticas de los Pacientes

Pantalla de Estadísticas de las Enfermedades Asociadas

En esta pantalla se observa las estadísticas de hipertensión arterial sistémica, de diabetes mellitus tipo 2 y del tratamiento antihipertensivo del conjunto total de registros de los pacientes seleccionados del Hospital Ceibos.



Fig. 15 Pantalla de Estadísticas de Enfermedades Asociadas

Pantalla de inserción de variables para el perfil lipídico

En esta pantalla se observa la estructura del perfil lipídico del paciente a analizar, que consta del colesterol total, el sexo del paciente, la lipoproteína de alta densidad (hdl-c), el colesterol no hdl-c, la lipoproteína de baja densidad (ldl-c) y los triglicéridos.

The screenshot shows a web interface for calculating a lipid profile. The page title is 'CARDIO-ZT' and the breadcrumb is 'Modelos / Home / Perfil Lipídico'. The main heading is 'Calcular Perfil Lipídico'. The form contains the following fields and values:

Variable	Valor
Colesterol total:	400
Sexo:	Hombre
HDL-C:	70
NO-HDL-C:	330
LDL-C:	128
Triglicéridos:	325

Below the input fields is a blue 'Calcular' button. Underneath, the 'Resultados:' section is currently empty, displaying 'Nada que mostrar'. At the bottom, it indicates 'Updated yesterday at 11:59 PM'.

Fig. 16 Pantalla de inserción de variables para el perfil lipídico

Pantalla del nivel del perfil lipídico

En esta pantalla se observa el resultado del cálculo de los niveles del perfil lipídico del paciente analizado en bandas de colores: color verde para estado satisfactorio y en color rojo para un estado no satisfactorio.

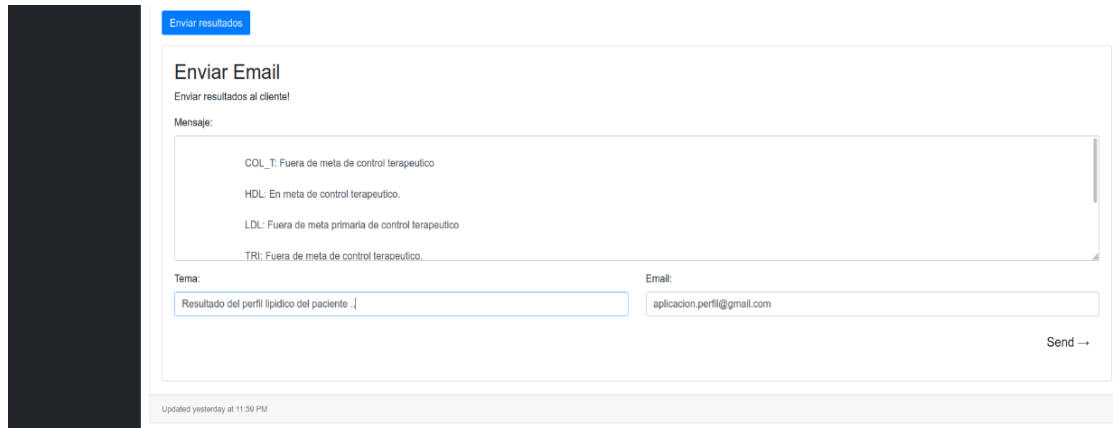
The screenshot shows the 'Resultados:' section with five horizontal bars indicating the status of each lipid component:

Variable	Estado
COL_T: Fuera de meta de control terapéutico	No satisfactorio (Rojo)
HDL: En meta de control terapéutico.	Satisfactorio (Verde)
LDL: Fuera de meta primaria de control terapéutico	No satisfactorio (Rojo)
TR: Fuera de meta de control terapéutico.	No satisfactorio (Rojo)
NO_HDL: En meta secundaria de control terapéutico.	Satisfactorio (Verde)

Fig. 17 Pantalla del nivel del perfil lipídico

Pantalla de envío de resultados del perfil lipídico

En esta pantalla se visualiza el envío de los resultados de los niveles del perfil lipídico del paciente o la paciente seleccionada para el médico especialista en medicina interna y tener así una interacción médico-paciente del estado de su salud.



The screenshot shows a web interface for sending an email. At the top left, there is a blue button labeled "Enviar resultados". Below it, the main heading is "Enviar Email" with the subtitle "Enviar resultados al cliente!". The "Mensaje:" field contains the following text: "COL_T: Fuera de meta de control terapeutico", "HDL: En meta de control terapeutico.", "LDL: Fuera de meta primaria de control terapeutico", and "TRI: Fuera de meta de control terapeutico.". The "Tema:" field is filled with "Resultado del perfil lipidico del paciente .!". The "Email:" field is filled with "aplicacion.perfil@gmail.com". A "Send -->" button is located at the bottom right. At the bottom left, there is a small text "Updated yesterday at 11:59 PM".

Fig. 18 Pantalla de envío de resultados del perfil lipídico

Anexo 14: Certificación y traducción del resumen.



Universidad
Nacional
de Loja

FACULTAD E LA EDUCACIÓN, EL ARTE Y LA COMUNICACIÓN
CARRERA DE PEDAGOGÍA DE LOS IDIOMA NACIONALES Y EXTRANJEROS

Mgs. Miriam Eucevia Troya Sánchez

**DOCENTE DE LA CARRERA DE PEDAGOGÍA DE LOS IDIOMAS NACIONALES
Y EXTRANJEROS DE LA UNIVERSIDAD NACIONAL DE LOJA**

C E R T I F I C A:

Haber realizado la traducción de Español – Inglés del resumen del Trabajo de Titulación previo a la obtención del título de Ingeniero en Sistemas, titulado **APLICACIÓN WEB PARA DETERMINAR EL PERFIL LIPÍDICO Y SU RELACIÓN CON EL RIESGO CARDIOVASCULAR EN PACIENTES DIABÉTICOS TIPO 2 E HIPERTENSOS DE LA CONSULTA EXTERNA DEL HOSPITAL CEIBOS** de autoría del estudiante **Roger Alexander Torres Yaguana** con C. I: 1103491435. Se autoriza al interesado hacer uso de la misma para los trámites que crea conveniente. Es todo cuanto puedo certificar en honor a la verdad.

Emitida en Loja, a 1 día del mes de Agosto de 2022.



MIRIAM EUCEVIA
TROYA SANCHEZ

Mgs. Miriam Eucevia Troya Sánchez
Licenciada en Ciencias de la Educación. Mención Inglés
REGISTRO EN LA SENECYT N° 1008-02- 154730