



Universidad
Nacional
de Loja

Universidad Nacional de Loja

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

Carrera de Electromecánica

**Predicción de fallas en los aerogeneradores de la Central Eólica Villonaco,
aplicando la técnica de aprendizaje profundo basado en transformadores.**

**Trabajo de Integración Curricular,
previo a la obtención del título de
Ingeniero Electromecánico.**

AUTOR:

Joel Stiven Torres Cabrera

DIRECTOR:

Ing. Jorge Luis Maldonado Correa

Loja – Ecuador

2024

Certificación

Loja, 29 de abril del 2024

Ing. Jorge Luis Maldonado Correa. Mg. Sc.

DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR

CERTIFICO:

Que he revisado y orientado todo proceso de la elaboración del Trabajo de Integración Curricular denominado: **Predicción de fallas en los aerogeneradores de la Central Eólica Villonaco, aplicando la técnica de aprendizaje profundo basado en transformadores**, de autoría del estudiante **Joel Stiven Torres Cabrera**, con **cédula de identidad Nro. 1150343463**, previo a la obtención del título de **Ingeniero Electromecánico**, una vez que el trabajo cumple con todos los requisitos exigidos por la Universidad Nacional de Loja para el efecto, autorizo la presentación para la respectiva sustentación y defensa.

Ing. Jorge Luis Maldonado Correa. Mg. Sc.

DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR

Autoría

Yo, **Joel Stiven Torres Cabrera**, declaro ser autor del presente Trabajo de Integración Curricular y eximo expresamente a la Universidad Nacional de Loja y a sus representantes jurídicos de posibles reclamos y acciones legales, por el contenido del mismo. Adicionalmente, acepto y autorizo a la Universidad Nacional de Loja la publicación de mí del Trabajo de Integración Curricular o de titulación en el Repositorio Digital Institucional – Biblioteca Virtual.

Firma

Cédula de identidad: 1150343463

Fecha: 29/04/2024

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

Teléfono: 0979701823

Carta de autorización por parte del autor, para la consulta, de producción parcial o total y/o publicación electrónica de texto completo, del Trabajo de Integración Curricular.

Yo, **Joel Stiven Torres Cabrera**, declaro ser autor del Trabajo de Integración Curricular denominado: **Predicción de fallas en los aerogeneradores de la Central Eólica Villonaco, aplicando la técnica de aprendizaje profundo basado en transformadores**, como requisito para optar por el título de **Ingeniero Electromecánico**, autorizo al sistema Bibliotecario de la Universidad Nacional de Loja para que, con fines académicos, muestre la producción intelectual de la Universidad, a través de la visibilidad de su contenido 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 el plagio o copia del Trabajo de Integración Curricular que realice un tercero.

Para constancia de esta autorización, en la ciudad de Loja, a los veintinueve días de abril del dos mil veinticuatro.

Firma

Autor: Joel Stiven Torres Cabrera

Cédula: 1150343463

Dirección: Loja, (Barrio Julio Ordoñez)

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

Teléfono: 0979701823

DATOS COMPLEMENTARIOS:

Director del Trabajo de Integración Curricular: Ing. Jorge Luis Maldonado Correa

Dedicatoria

El presente trabajo de investigación está dedicado a mis padres, Gustavo Torres y Laura Cabrera, quienes con su trabajo, amor, consejos y paciencia me han apoyado durante mi formación académica y me han ayudado a lograr este objetivo.

A mis hermanos José, Yonny, Marlon y María por ser parte importante de mi vida, su apoyo y valiosos consejos para seguir avanzando en mis estudios y mejorando cada día.

A mis abuelos, Ilduara y Gerisneldo, quienes me brindaron su apoyo y cobijo cuando más lo necesité. Siempre les estaré agradecido.

A mis profesores, por su formación y conocimientos impartidos a lo largo de mi trayectoria académica.

Finalmente, dedico este trabajo a mi familia, a mi novia y amigos con quienes he compartido gratos momentos y quienes han sido cruciales para ayudarme a alcanzar mis metas.

Joel Stiven Torres Cabrera

Agradecimiento

En primer lugar, agradezco profundamente a mis padres quienes me han brindado sus enseñanzas, consejos y sacrificio para seguir adelante; a mis hermanos por su motivación y apoyo incondicional, así mismo a Dios por brindarme la salud y la fortaleza necesaria para alcanzar este hito tan importante.

A toda mi familia y amigos que han estado a mi lado durante todos estos años de formación académica, apoyándome para cumplir con este objetivo.

A Xavier y Albin quienes estuvieron presentes y compartieron gratas experiencias que me ayudaron a lograr este objetivo.

A mi novia, Karla, por ser mi fuente constante de motivación y apoyo durante mi travesía universitaria, ello ha sido fundamental para mantenerme firme en el camino hacia mis metas académicas y profesionales.

A mi director de trabajo de integración curricular, Ing. Jorge Maldonado, por su experiencia, asesoría, consejos y guía hasta culminar con éxito este proyecto de investigación.

A la Universidad Nacional de Loja, a la Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, y a la carrera de Ingeniería Electromecánica, a cada uno de mis profesores y autoridades, quienes me han inculcado sus conocimientos y experiencias para lograr formarme como un futuro profesional de la patria.

Finalmente, quisiera expresar mi agradecimiento al Centro de Investigaciones Tecnológicas y Energéticas (CITE) por brindarme la información necesaria para llevar a cabo este proyecto de investigación.

Joel Stiven Torres Cabrera

Índice de contenidos

Portada	i
Certificación	ii
Autoría	iii
Dedicatoria	v
Agradecimiento	vi
Índice de contenidos	vii
Índice de tablas	xv
Índice de figuras	xvii
Índice de anexos	xxix
Simbología	xxx
1. Título	1
2. Resumen	2
Abstract	3
3. Introducción	4
4. Marco teórico	6
4.1 Capítulo I: Energía Eólica	6
4.1.1 Central Eólica Villonaco	6
4.1.2 Aerogenerador	7
4.1.3 Componentes de un aerogenerador	7
4.1.3.1 Torre	8
4.1.3.2 Aspas	8
4.1.3.3 Sistema de orientación de palas (<i>Pitch control</i>)	8
4.1.3.4 Sistema de orientación de la góndola (<i>Yaw control</i>)	9
4.1.3.5 Góndola	9
4.1.3.6 Velea de viento	9
4.1.3.7 Controlador	9
4.1.3.8 Sistema eléctrico	9

4.1.3.9	Frenos	9
4.1.4	Sistema SCADA	9
4.1.5	Datos de series temporales	10
4.1.6	Mantenimiento en aerogeneradores	10
4.1.7	Tipos de mantenimiento	10
4.1.7.1	Mantenimiento correctivo	10
4.1.7.2	Mantenimiento preventivo	10
4.1.7.3	Mantenimiento predictivo	11
4.1.8	Fallas en aerogeneradores	11
4.1.8.1	Convertidor electrónico	12
4.1.8.2	Generador eléctrico	13
4.1.8.3	Sistema hidráulico	13
4.1.8.4	Sensores	14
4.1.8.5	Sistema de control	14
4.1.8.6	Sistema de orientación de palas.	14
4.2	Capítulo II: Procesamiento de datos	14
4.2.1	Preprocesamiento de datos	14
4.2.2	Integridad de los datos	16
4.2.2.1	Datos faltantes	16
4.2.2.1.1	Eliminación de filas dañinas	16
4.2.2.1.2	Relleno de valores perdidos	16
4.2.3	Ingeniería de características	16
4.2.3.1	Selección de características	16
4.2.3.1.1	Método de eliminación recursivo de características (RFE por sus siglas de Recursive Feature Elimination)	17
4.2.3.1.2	Método estadístico de correlación	17
4.2.3.1.3	Método estadístico de prueba de hipótesis	17
4.2.3.1.4	Métricas de selección basado en árboles	17
4.2.3.1.5	Modelos lineales y regularización	18
4.2.3.2	Escalado y normalización de características	18
4.2.4	Datos de entrenamiento, validación y prueba	18
4.3	Capítulo III: Estado del arte	19
4.3.1	Inteligencia artificial	19
4.3.2	Aprendizaje automático (Machine Learning)	19

4.3.2.1	Detección de fallas en aerogeneradores con aprendizaje automático	20
4.3.3	Aprendizaje profundo (Deep Learning)	21
4.3.4	Tipos de aprendizaje de aprendizaje profundo	21
4.3.4.1	Aprendizaje supervisado	21
4.3.4.2	Aprendizaje no supervisado	22
4.3.4.3	Aprendizaje por refuerzo	22
4.3.5	Detección de anomalías con aprendizaje profundo	22
4.3.6	Mecanismo de atención: El Modelo del Transformador (The Transformer Model)	22
4.3.6.1	Codificador	23
4.3.6.2	Decodificador	24
4.3.6.3	Atención	25
4.3.7	Detección de anomalías con la arquitectura de transformadores	26
4.3.8	Detección de fallas en aerogeneradores con arquitectura de transformadores	28
4.3.9	Métodos de optimización modernos	28
4.3.10	Métricas de evaluación	29
4.3.10.1	Matriz de confusión en predicción	29
4.3.10.2	Precisión (Precision)	29
4.3.10.3	Sensibilidad (Recall) o Tasa de verdadero positivo	30
4.3.10.4	F1 Score	30
4.3.11	Problemas típicos del aprendizaje automático	30
4.3.11.1	Calidad de los datos	30
4.3.11.2	Dimensionalidad	30
4.3.11.3	Subajuste (Underfitting)	30
4.3.11.4	Sobreajuste (Overfitting)	30
5.	Metodología	31
5.1	Área de estudio	31
5.2	Características de los aerogeneradores en estudio.	31
5.3	Materiales y equipos	36
5.3.1	Material bibliográfico	36
5.3.2	Equipos Tecnológicos	37
5.3.3	Recursos humanos	37
5.3.4	Lenguaje de programación	37

5.3.4.1	Librerías	37
5.3.5	Acuerdo de confidencialidad	37
5.4	Procedimiento	37
5.4.1	Enfoque metodológico	37
5.4.2	Tipo de diseño	37
5.5	Metodología	37
5.6	Procedimiento y análisis de datos del objetivo 1	38
5.6.1	Procesamiento de datos	38
5.6.1.1	Análisis de datos del sistema SCADA y procedimiento	38
5.6.1.2	Procesamiento de valores faltantes y perdidos.	38
5.6.2	Filtrado de información del conjunto de datos	41
5.6.3	Etiquetado de datos	41
5.6.4	Selección de características	43
5.6.4.1	Escalado y normalización	43
5.6.4.2	División de los datos de entrenamiento, validación y prueba	43
5.6.4.3	Análisis y comportamiento de variables SCADA	43
5.6.5	Desarrollo y aplicación del modelo	44
5.6.5.1	Implementación del primer modelo del Transformador Anomaly Transformer	45
5.6.5.1.1	Atención de anomalías	46
5.6.5.1.2	Discrepancia de asociación	48
5.6.5.1.3	Aprendizaje de asociación min-max	48
5.6.5.1.4	Estrategia de mini-max	49
5.6.5.1.5	Criterios de anomalía basados en relaciones	50
5.6.5.2	Implementación del segundo modelo del Transformador TranAD	50
5.6.5.2.1	Entrenamiento adversarial de dos fases fuera de línea	53
5.6.5.2.1	Detección de anomalías	56
5.6.6	Entrenamiento del modelo	57
5.6.7	Procedimiento de evaluación del modelo	57
6.	Resultados	58
6.1	Hardware	58
6.2	Software	58
6.3	Procesamiento de datos	58

6.4	Integridad de los datos	59
6.5	Filtrado de información del conjunto de datos	59
6.6	Etiquetado de datos	60
6.7	Análisis estadístico	61
6.7.1	Descripción de las variables de estudio	61
6.7.2	Curva de potencia y alarmas	62
6.7.3	Curva de Potencia vs Temperatura IGBT	64
6.7.4	Relación de las variables a través del tiempo	64
6.7.5	Gráfico de control	66
6.7.6	Gráfico de caja y bigote	66
6.8	Selección de características	67
6.8.1	Método de eliminación recursivo de características (RFE).	67
6.8.2	Método estadístico de correlación (C)	67
6.8.3	Método estadístico de prueba de hipótesis (PV)	68
6.8.4	Métricas de selección basado en árboles (AD)	69
6.8.5	Modelos lineales para selección de características (ML)	69
6.8.6	Escalado de características	69
6.8.7	División de los datos	70
6.9	Aplicación del primer modelo de detección de anomalías (Anomaly Transformer)	70
6.9.1	Tamaño de ventana deslizante	70
6.9.2	Ventana deslizante para detección de fallas (primera etapa)	71
6.9.3	Evaluación de detección de fallas	71
6.9.4	Detección de fallas en WT1 mediante Anomaly Transformer	71
6.9.4.1	Anomaly Transformer utilizando el método RFE1 en WT1	72
6.9.4.2	Anomaly Transformer utilizando el método RFE12 en WT1	72
6.9.4.3	Anomaly Transformer utilizando el método RF24 en WT1	73
6.9.4.4	Anomaly Transformer utilizando el método C6 en WT1	74
6.9.4.5	Anomaly Transformer utilizando el método PV18 en WT1	74
6.9.4.6	Anomaly Transformer utilizando el método AD23 en WT1	75
6.9.4.7	Anomaly Transformer utilizando el método LM18 en WT1	75
6.9.4.8	Resultados de Anomaly Transformer para WT1	76
6.9.5	Detección de fallas en WT2 mediante Anomaly Transformer	76
6.9.5.1	Anomaly Transformer utilizando el método RFE 12 en WT2	77
6.9.5.2	Resultados de Anomaly Transformer para WT2	77

6.9.6	Detección de fallas en WT3 mediante Anomaly Transformer	77
6.9.6.1	Anomaly Transformer utilizando el método RFE 12 en WT3	77
6.9.6.2	Resultados de Anomaly Transformer para WT3	78
6.9.7	Detección de fallas en WT4 mediante Anomaly Transformer	78
6.9.7.1	Anomaly Transformer utilizando el método LM18 en WT4	78
6.9.7.2	Resultados de Anomaly Transformer para WT4	79
6.9.8	Detección de fallas en WT5 mediante Anomaly Transformer	79
6.9.8.1	Anomaly Transformer utilizando el método RFE 12 en WT5	79
6.9.8.2	Resultados de Anomaly Transformer para WT5	80
6.9.9	Detección de fallas en WT6 mediante Anomaly Transformer	80
6.9.9.1	Anomaly Transformer utilizando el método AD23 en WT6	80
6.9.9.2	Resultados de Anomaly Transformer para WT6	81
6.9.10	Detección de fallas en WT7 mediante Anomaly Transformer	81
6.9.10.1	Anomaly Transformer utilizando el método RFE 12 en WT7	81
6.9.10.2	Resultados de Anomaly Transformer para WT7	82
6.9.11	Detección de fallas en WT8 mediante Anomaly Transformer	82
6.9.11.1	Anomaly Transformer utilizando el método C6 en WT8	82
6.9.11.2	Resultados de Anomaly Transformer para WT8	83
6.9.12	Detección de fallas en WT9 mediante Anomaly Transformer	83
6.9.12.1	Anomaly Transformer utilizando el método AD23 en WT9	83
6.9.12.2	Resultados de Anomaly Transformer para WT9	84
6.9.13	Detección de fallas en WT10 mediante Anomaly Transformer	84
6.9.13.1	Anomaly Transformer utilizando el método C6 en WT10	84
6.9.13.2	Resultados de Anomaly Transformer para WT10	85
6.9.14	Detección de fallas en WT11 mediante Anomaly Transformer	85
6.9.14.1	Anomaly Transformer utilizando el método C6 en WT11	85
6.9.14.2	Resultados de Anomaly Transformer para WT11	86
6.10	Aplicación del segundo modelo de detección de anomalías TranAD	87
6.10.1	Ventana deslizante para detección de fallas (segunda etapa)	87
6.10.2	Detección de fallas en WT1 mediante TranAD	87
6.10.2.1	TranAD utilizando el método RF12 en WT1	87
6.10.2.2	TranAD utilizando el método C6 en WT1	89
6.10.2.3	TranAD utilizando el método PV18 en WT1	90
6.10.2.4	TranAD utilizando el método LM18 en WT1	90

6.10.2.5	Resultados de TranAD para WT1	91
6.10.3	Detección de fallas en WT2 mediante TranAD	91
6.10.3.1	TranAD utilizando el método PV18 en WT2	91
6.10.3.2	Resultados de TranAD para WT2	92
6.10.4	Detección de fallas en WT3 mediante TranAD	92
6.10.4.1	TranAD utilizando el método PV18 en WT3	92
6.10.4.2	Resultados de TranAD para WT3	93
6.10.5	Detección de fallas en WT4 mediante TranAD	93
6.10.5.1	TranAD utilizando el método PV18 en WT4	93
6.10.5.2	Resultados de TranAD para WT4	94
6.10.6	Detección de fallas en WT5 mediante TranAD	94
6.10.6.1	TranAD utilizando el método RFE12 en WT5	94
6.10.6.2	Resultados de TranAD para WT5	95
6.10.7	Detección de fallas en WT6 mediante TranAD	95
6.10.7.1	TranAD utilizando el método C6 en WT6	95
6.10.7.2	Resultados de TranAD para WT6	95
6.10.8	Detección de fallas en WT7 mediante TranAD	96
6.10.8.1	TranAD utilizando el método LM18 en WT7	96
6.10.8.2	Resultados de TranAD para WT7	96
6.10.9	Detección de fallas en WT8 mediante TranAD	96
6.10.9.1	TranAD utilizando el método PV18 en WT8	96
6.10.9.2	Resultados de TranAD para WT8	97
6.10.10	Detección de fallas en WT9 mediante TranAD	97
6.10.10.1	TranAD utilizando el método RFE12 en WT9	97
6.10.10.2	Resultados de TranAD para WT9	98
6.10.11	Detección de fallas en WT10 mediante TranAD	98
6.10.11.1	TranAD utilizando el método RFE12 en WT10	98
6.10.11.2	Resultados de TranAD para WT10	98
6.10.12	Detección de fallas en WT11 mediante TranAD	99
6.10.12.1	TranAD utilizando el método PV18 en WT11	99
6.10.12.2	Resultados de TranAD para WT11	99
6.11	Resultados Generales	100

7. Discusión **101**

8. Conclusiones	103
9. Recomendaciones	105
10. Bibliografía	106
11. Anexos	112

Índice de tablas:

Tabla 1. Resumen de fallas, señales y métodos utilizados para procesar datos.	15
Tabla 2. Matriz de Confusión en predicción.	29
Tabla 3. Datos técnicos del aerogenerador GOLDWIND GW 70/1 500.	31
Tabla 4. Formato de archivo de registro de operación y mantenimiento de la CEV.	41
Tabla 5. Archivo Excel de alarmas dado por el sistema SCADA.	41
Tabla 6. Desafíos encontrados en los archivos del SCADA.	58
Tabla 7. Datos del sistema SCADA.	59
Tabla 8. Variables eliminadas que no serán parte de este proceso.	59
Tabla 9. Filtrado de fallas y código de alarmas.	60
Tabla 10. Datos SCADA, alarmas y fallas del módulo IGBT.	61
Tabla 11. Etiquetado de fallas de los datos SCADA.	61
Tabla 12. Descripción de valores de tendencia central.	61
Tabla 13. Resumen de las variables seleccionadas por diferentes métodos.	69
Tabla 14. División de los datos para los 11 aerogeneradores.	70
Tabla 15. Resultados de rendimiento del modelo AT para WT1.	76
Tabla 16. Resultados de rendimiento del modelo AT para WT2.	77
Tabla 17. Resultados de rendimiento del modelo AT para WT3.	78
Tabla 18. Resultados de rendimiento del modelo AT para WT4.	79
Tabla 19. Resultados de rendimiento del modelo AT para WT5.	80
Tabla 20. Resultados de rendimiento del modelo AT para WT6.	81
Tabla 21. Resultados de rendimiento del modelo AT para WT7.	82
Tabla 22. Resultados de rendimiento del modelo AT para WT8.	83
Tabla 23. Resultados de rendimiento del modelo AT para WT9.	84
Tabla 24. Resultados de rendimiento del modelo AT para WT10.	85
Tabla 25. Resultados de rendimiento del modelo AT para WT11.	86
Tabla 26. Resultados de rendimiento del modelo TranAD para WT1.	91
Tabla 27. Resultados de rendimiento del modelo TranAD para WT3.	92
Tabla 28. Resultados de rendimiento del modelo TranAD para WT3.	93
Tabla 29. Resultados de rendimiento del modelo TranAD para WT4.	94
Tabla 30. Resultados de rendimiento del modelo TranAD para WT5.	95
Tabla 31. Resultados de rendimiento del modelo TranAD para WT6.	95
Tabla 32. Resultados de rendimiento del modelo TranAD para WT7.	96

Tabla 33. Resultados de rendimiento del modelo TranAD para WT8.....	97
Tabla 34. Resultados de rendimiento del modelo TranAD para WT9.....	98
Tabla 35. Resultados de rendimiento del modelo TranAD para WT10.....	98
Tabla 36. Resultados de rendimiento del modelo TranAD para WT11.....	99
Tabla 37. Resultados generales del rendimiento de Anomaly Transformer.	100
Tabla 38. Resultados generales del rendimiento de TranAD.....	100
Tabla 39. Anexo de descripción de variables SCADA.	112
Tabla 40. Descripción de códigos de falla del sistema SCADA.....	114

Índice de figuras:

Figura 1. Central Eólica Villonaco.....	7
Figura 2. Componentes de un aerogenerador GW70/1500.....	8
Figura 3. Porcentaje de fallas en los componentes de los aerogeneradores onshore.....	12
Figura 4. Fallas en los componentes del convertidor electrónico de potencia.....	13
Figura 5. Modelo general de aprendizaje en un modelo de aprendizaje automático.....	20
Figura 6. Arquitectura del transformador.....	24
Figura 7. Metodología utilizada en la elaboración del proyecto.....	38
Figura 8. Procesado de datos.....	40
Figura 9. Flujograma de etiquetado de fallas y alarmas del módulo IGBT.....	42
Figura 10. Flujograma de modelo de predicción de fallas.....	44
Figura 11. Metodología de Anomaly Transformer.....	46
Figura 12. Metodología de TranAD.....	51
Figura 13. Curva de potencia WT1.....	62
Figura 14. Curva de potencia para todos los aerogeneradores.....	63
Figura 15. Curva de Potencia vs Temperatura IGBT.....	64
Figura 16. Serie temporal de variables relacionadas con la falla del módulo IGBT en WT1.....	65
Figura 17. Serie temporal de la temperatura del IGBT.....	66
Figura 18. Diagrama de caja y bigotes de las variables del sistema SCADA.....	67
Figura 19. Matriz de correlación de las variables del sistema SCADA.....	68
Figura 20. Resultados para RFE1 en el análisis de temperatura del módulo IGBT mediante el modelo AT para WT1.....	72
Figura 21. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.....	73
Figura 22. Resultados del modelo para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.....	74
Figura 23. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.....	74
Figura 24. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.....	75
Figura 25. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.....	75

Figura 26. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.	76
Figura 27. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.	77
Figura 28. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	78
Figura 29. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante modelo AT para WT4.	79
Figura 30. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.	80
Figura 31. Resultados para AD23 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	81
Figura 32. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	82
Figura 33. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.	83
Figura 34. Resultados para AD23 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.	84
Figura 35. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante modelo AT para WT10.	85
Figura 36. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	86
Figura 37. Resultados para RFE12 en el análisis de evaluación de primera etapa mediante el modelo TranAD para WT1.	88
Figura 38. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.	89
Figura 39. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.	89
Figura 40. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.	90
Figura 41. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.	90
Figura 42. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.	92

Figura 43. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.	93
Figura 44. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.	94
Figura 45. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.	94
Figura 46. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.	95
Figura 47. Resultados para LM18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.	96
Figura 48. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.	97
Figura 49. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.	97
Figura 50. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.	98
Figura 51. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.	99
Figura 52. Resultados para RFE1 en el análisis fallas en el módulo IGBT mediante el modelo AT para WT2.	116
Figura 53. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.	116
Figura 54. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.	117
Figura 55. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.	117
Figura 56. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.	118
Figura 57. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.	118
Figura 58. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.	119
Figura 59. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	119

Figura 60. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	120
Figura 61. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	120
Figura 62. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	121
Figura 63. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	121
Figura 64. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	122
Figura 65. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.	122
Figura 66. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	123
Figura 67. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	123
Figura 68. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	124
Figura 69. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	124
Figura 70. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	125
Figura 71. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	125
Figura 72. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	126
Figura 73. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.	126
Figura 74. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.	127
Figura 75. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.	127
Figura 76. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.	128

Figura 77. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.	128
Figura 78. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.	129
Figura 79. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.	129
Figura 80. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	130
Figura 81. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	130
Figura 82. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	131
Figura 83. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	131
Figura 84. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	132
Figura 85. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	132
Figura 86. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.	133
Figura 87. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	133
Figura 88. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	134
Figura 89. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	134
Figura 90. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	135
Figura 91. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	135
Figura 92. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	136
Figura 93. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.	136

Figura 94. Resultados para RFE1 en el análisis de temperatura del módulo IGBT mediante el modelo AT para WT8.	137
Figura 95. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.	137
Figura 96. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.	138
Figura 97. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.	138
Figura 98. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.	139
Figura 99. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.	139
Figura 100. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.	140
Figura 101. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.	140
Figura 102. Resultados para RFE12 en el análisis de fallas del módulo mediante el modelo AT para WT9.	141
Figura 103. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.	141
Figura 104. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.	142
Figura 105. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.	142
Figura 106. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.	143
Figura 107. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.	143
Figura 108. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.	144
Figura 109. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.	144
Figura 110. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.	145

Figura 111. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.	145
Figura 112. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.	146
Figura 113. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.	146
Figura 114. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.	147
Figura 115. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	147
Figura 116. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	148
Figura 117. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	148
Figura 118. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	149
Figura 119. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	149
Figura 120. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	150
Figura 121. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.	150
Figura 122. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT1.	151
Figura 123. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.	152
Figura 124. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT1.....	152
Figura 125. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.	153
Figura 126. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT1.....	153
Figura 127. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.	154

Figura 128. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.	154
Figura 129. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.	155
Figura 130. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.	155
Figura 131. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.	156
Figura 132. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.....	156
Figura 133. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.	157
Figura 134. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.....	157
Figura 135. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.	158
Figura 136. Resultados para RFE12 en el análisis de advertencias mediante el modelo TranAD para WT3.	158
Figura 137. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.	159
Figura 138. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT3.	159
Figura 139. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.	160
Figura 140. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT3.....	160
Figura 141. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.	161
Figura 142. Resultados para LM18 en el análisis de advertencias variable a variable en relación a las fallas en el módulo IGBT mediante el modelo TranAD para WT3.....	161
Figura 143. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.	162
Figura 144. Resultados para RFE12 en el análisis de advertencias variable mediante el modelo TranAD para WT4.	162

Figura 145. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.	163
Figura 146. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT4.	163
Figura 147. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.	164
Figura 148. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT4.....	164
Figura 149. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.	165
Figura 150. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT4.....	165
Figura 151. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.	166
Figura 152. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para la WT5.	166
Figura 153. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.	167
Figura 154. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT5.	167
Figura 155. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.	168
Figura 156. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT5.....	168
Figura 157. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.	169
Figura 158. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT5.....	169
Figura 159. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.	170
Figura 160. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.	170
Figura 161. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.	171

Figura 162. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.	171
Figura 163. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.	172
Figura 164. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.....	172
Figura 165. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.	173
Figura 166. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.....	173
Figura 167. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.	174
Figura 168. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT7.	174
Figura 169. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.	175
Figura 170. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT7.	175
Figura 171. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.	176
Figura 172. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT7.....	176
Figura 173. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.	177
Figura 174. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para la WT7.....	177
Figura 175. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.	178
Figura 176. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.	178
Figura 177. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.	179
Figura 178. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.	179

Figura 179. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.	180
Figura 180. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.....	180
Figura 181. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.	181
Figura 182. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.....	181
Figura 183. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.	182
Figura 184. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.....	182
Figura 185. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.	183
Figura 186. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.	183
Figura 187. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.	184
Figura 188. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.....	184
Figura 189. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.	185
Figura 190. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.....	185
Figura 191. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.	186
Figura 192. Resultados para RFE12 en el análisis de advertencias en el módulo IGBT mediante el modelo TranAD para WT10.....	186
Figura 193. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.	187
Figura 194. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT10.	187
Figura 195. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.	188

Figura 196. Resultados para PV18 en el análisis advertencias variable a variable mediante el modelo TranAD para WT10.	188
Figura 197. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.	189
Figura 198. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT10.....	189
Figura 199. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.	190
Figura 200. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT11.....	190
Figura 201. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.	191
Figura 202. Resultados para C6 en el análisis de advertencias en el módulo IGBT mediante el modelo TranAD para WT11.	191
Figura 203. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.	192
Figura 204. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT11.....	192
Figura 205. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.	193
Figura 206. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT11.....	193
Figura 207. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.	194
Figura 208. Archivos de dataset de operación de los aerogeneradores.....	195
Figura 209. Archivos de alarmas del SCADA y de mantenimiento de los aerogeneradores.	196
Figura 210. Celda de código de ruta de archivos para procesamiento de datos.....	196
Figura 211. Curva de potencia del aerogenerador GOLDWIND GW/1500.....	199
Figura 212. Acuerdo de confidencialidad	200
Figura 213. Potencia vs Temperatura IGBT	201
Figura 214. Certificado de traducción del Resumen.....	202

Índice de anexos:

Anexo 1. Variables SCADA	112
Anexo 2. Códigos de falla y descripción de alarmas SCADA.....	114
Anexo 3. Resultados detallados para el primer modelo Anomaly Transformer	116
Anexo 4. Resultados para toda la selección de características desde la WT1 hasta la WT11 para TranAD.....	151
Anexo 5. Guía para ejecutar de manera efectiva los resultados de este trabajo	195
Anexo 6. Curva de potencia del aerogenerador GOLDWIND GW 70/1500	199
Anexo 7. Acuerdo de confidencialidad	200
Anexo 8. Potencia vs Temperatura IGBT	201
Anexo 9. Certificado de traducción del Resumen.....	202

Simbología:

Variable	Descripción
<i>IA</i>	Inteligencia Artificial
<i>O&M</i>	Operación y mantenimiento
<i>SCADA</i>	Supervisión, Control y Adquisición de Datos
<i>CEV</i>	Central Eólica Villonaco
<i>CBM</i>	Monitoreo basado en la condición.
<i>PCB</i>	Placa de circuito impreso.
<i>IGBT</i>	Transistor bipolar de puerta aislada.
<i>PCA</i>	Análisis de componentes principales
<i>PNL</i>	Procesamiento de lenguaje natural
<i>AST</i>	Transformador disperso adversario.
<i>MAML</i>	Meta Aprendizaje Independiente del Modelo
<i>GAN</i>	Redes neuronales generativas adversarias
<i>WT</i>	Aerogenerador

1. Título

Predicción de fallas en los aerogeneradores de la Central Eólica Villonaco, aplicando la técnica de aprendizaje profundo basado en transformadores.

2. Resumen

La energía eólica representa una alternativa renovable y sostenible que puede contribuir a reducir la contaminación generada los combustibles fósiles. La Central Eólica Villonaco (CEV) abastece al sistema eléctrico ecuatoriano, y una parada ocasionada por una falla en sus aerogeneradores puede ser costosa e implicar un largo tiempo de inactividad. Por ello, es importante predecir fallas en los diferentes componentes del sistema. Esta investigación se enfoca en el convertidor, especialmente en el módulo IGBT, debido a que este componente es el que más reemplazos ha tenido durante el tiempo de funcionamiento de la CEV. El aprendizaje profundo basado en el modelo de transformadores es una técnica nueva, potente e ideal para trabajar en un gran conjunto de datos. Por lo tanto, se utilizará para predecir fallas del módulo IGBT en 11 aerogeneradores de la CEV. El dataset de funcionamiento de los aerogeneradores contiene 1,63 GB de datos; además, se tiene el archivo de alarmas del SCADA (Supervisory Control and Data Acquisition) y el plan de Operación y Mantenimiento (O&M) durante el periodo 2014-2021. Estos datos se preprocesan y se aplica métodos de selección de variables; además, se realiza un proceso de filtrado de datos para garantizar que los datos correspondan únicamente al funcionamiento “normal” del aerogenerador. Luego, se aplica dos metodologías basada en Transformadores, capaz de detectar anomalías en datos SCADA. A estas metodologías, se implementó una evaluación de 2 etapas, para intentar predecir la ocurrencia de una falla. Los resultados experimentales muestran que el modelo es capaz de predecir fallas en el módulo IGBT con un F1 Score promedio superior al 90% y un avance promedio de 5.08 meses.

***Palabras clave:** Aerogeneradores, Análisis de datos SCADA, Módulo IGBT, aprendizaje profundo, Transformadores, detección de anomalías profundas.*

Abstract

Wind energy represents a renewable and sustainable alternative that can contribute to reducing pollution generated by the use of fossil fuels. The Villonaco wind farm (CEV) supplies the Ecuadorian electrical system, and a stoppage caused by a failure in its wind turbines can be costly and involve long downtime. Therefore, it is important to predict failures in the different components. This research focuses on the converter, especially on the IGBT module, because this component has had the most replacements during the operation time of the CEV. Deep learning based on transformer models is a new and powerful technique that is ideal for working with large datasets. Therefore, it will be used to predict failures in the IGBT module in the 11 wind turbines of the CEV. Initially, there are 1.63 GB of data corresponding to the operation of the wind turbines, as well as the SCADA (Supervisory Control and Data Acquisition) alarm file and the Operation and Maintenance (O&M) plan carried out in the period 2014-2021. These data are preprocessed and variable selection methods are applied; in addition, a data filtering process is performed to ensure that the data correspond only to the “normal” operation of the wind turbine. Then, two methodologies based on Transformers, capable of detecting anomalies in SCADA data, are applied. To these methodologies, a 2-stage evaluation was implemented to try to predict the occurrence of a fault. Experimental results show that the model is able to predict failures in the IGBT module with an average F1 Score higher than 90% and an average advance of 5.08 months.

Keywords: *Wind turbines, SCADA data analysis, IGBT module, deep learning, Transformers, deep anomaly detection.*

3. Introducción

Ecuador es un país lleno de recursos naturales, y durante décadas los ha aprovechado para producir electricidad, por ello, en 2013 entraron varios proyectos emblemáticos en operación que han permitido un cambio de la matriz energética (Cevallos-Sierra & Ramos-Martin, 2018). Las centrales importantes de Ecuador son: Coca Codo Sinclair, Delsitanisagua, Minas San Francisco, Manduriacu, Mazar Dudas, Quijos, Sopladora y Toachi Pilatón, que representan una potencia total instalada de 2589,5 MW. Entre ellas se encuentra la Central Eólica Villonaco (CEV).

La CEV es el primer parque eólico de Ecuador continental, está ubicada en la provincia de Loja al sur del Ecuador, posee una potencia instalada de 16,5 MW, emplazado en un terreno montañoso a aproximadamente a 2720 m.s.n.m de coordenadas UTM 693030 E y 9556476 N. Entró en operación el año 2013, y consta de 11 aerogeneradores GOLDWIND GW70 clase “S”, con tecnología *Direct Drive* con generadores síncronos de 1,5 MW de potencia nominal (Maldonado-Correa et al., 2021).

La energía eólica es una de las fuentes de energía limpia de más rápido crecimiento en el mundo, pero depende en gran medida de los recursos de energía, Operación y Mantenimientos (O&Ms *abreviatura del inglés*) efectivos para generar suficiente energía, reducir tiempos de inactividad y costos de las turbinas eólicas (Chatterjee & Dethlefs, 2021). Es importante recalcar que en el Ecuador la CEV, es la primera experiencia a gran escala en este tipo de aprovechamiento de recurso energético por lo que supone desafíos importantes para la Empresa Pública Corporación Eléctrica del Ecuador CELEC.EP, encargada de su O&M.

La O&M representa una gran parte del costo total de las centrales eólicas en el mundo, y puede incrementar significativamente si existen fallas recurrentes en los aerogeneradores, por lo que establecer una estrategia de mantenimiento predictivo para la detección temprana de fallos potenciales en los aerogeneradores es esencial para reducir costos de O&M y para toma de decisiones de los operadores de las Centrales Eólicas (Maldonado-Correa et al., 2020).

El problema descrito anteriormente no es nuevo en la CEV, por lo que es necesario contribuir con un modelo de detección temprana de fallas utilizando el historial de datos del sistema SCADA como entrada principal y mediante algoritmos de aprendizaje profundo basado en transformadores se permita detectar fallas incipientes y monitorear el estado de “salud” en los aerogeneradores de la CEV. Esto permitirá reducir tiempos de mantenimiento por fallas en las unidades de generación y disminuir los costos en reparación de componentes, por consecuencia aumentar la producción de energía.

Teniendo en cuenta el enfoque de investigación se planteó:

Objetivos de investigación

Objetivo General

- Implementar un modelo de predicción de fallas en los aerogeneradores de la Central Eólica Villonaco utilizando aprendizaje profundo y los datos del sistema SCADA.

Objetivos específicos

- Detectar relaciones de comportamiento y de integridad en los datos de series temporales SCADA de la Central Eólica Villonaco para el aprovechamiento eficiente de la información.
- Aplicar modelos de aprendizaje profundo basado en transformadores en los datos de series temporales para predecir fallas en los aerogeneradores de la Central Eólica Villonaco.
- Evaluar los modelos de predicción desarrollados, utilizando métricas de rendimiento y precisión.

Con el objeto de cumplir con los objetivos, este trabajo se divide en secciones; la primera sección denominada Marco teórico, abarca conceptos fundamentales de la industria eólica, sistema SCADA, fallas en aerogeneradores, procesamiento de datos, ingeniería de características, inteligencia artificial, aprendizaje automático, aprendizaje profundo, arquitectura de transformador como método de detección de anomalías y métricas de evaluación. La segunda sección, Metodología, describe el procedimiento ejecutado, el procesamiento de los datos, los algoritmos de aprendizaje profundo basado en transformadores para detección de anomalías (Anomaly Transformer y TranAD) y evaluación de rendimiento de los modelos desarrollados. Por último, las secciones de Resultados, Conclusiones y Recomendaciones, en donde se detallan el procesamiento de los datos, el comportamiento de los datos, selección de variables y los resultados obtenidos de los modelos desarrollados de predicción de fallas en los 11 aerogeneradores de la CEV.

4. Marco teórico

4.1 Capítulo I: Energía Eólica

Los cambios en el sector eléctrico ecuatoriano han dependido en gran medida de los precios del petróleo, debido a que este ha sido la principal fuente de ingresos del país. Desde 2007, Ecuador ha experimentado grandes cambios debido al desarrollo de nuevas centrales hidroeléctricas y otras centrales de energías no convencionales por lo que la matriz energética ha estado cambiando de modo constante. Este cambio se debió a la necesidad de reducir la dependencia a los combustibles fósiles (Ponce-Jara et al., 2018).

Históricamente, el Ecuador no es un país que ha aprovechado sus recursos naturales, como el solar y el viento; según varios expertos, las condiciones topográficas y meteorológicas del país no son aptas para la implementación de energías renovables; sin embargo, las centrales eléctricas puestas en funcionamiento en las últimas décadas demuestran un factor de planta considerablemente alto, al realizar una comparación con otras centrales en el mundo (Cevallos-Sierra & Ramos-Martin, 2018).

La provincia de Loja tiene una población total de 511 184 habitantes y cubre un área de 11 026 km². Representa el 4% del total del país y es una zona donde los recursos son particularmente bendecidos por la naturaleza. Gracias a investigaciones previas, se ha descubierto que la provincia de Loja tiene un gran potencial en energía eólica y solar, lo que la convierte en una región de gran interés para el desarrollo sostenible.

Entre las centrales solares instaladas se tiene: Gonzaenergy, Lojaenergy, Renova Loja, Sabiango Solar y San Pedro, las cuales tienen una capacidad de generación instalada de 0,99 MW cada una. Además, cuenta con la CEV que tiene una potencia instalada de 16,5 MW que contribuye significativamente a la generación de energía eléctrica en el país.

Yang et al. (2018), mencionan que la energía eólica en la actualidad se está convirtiendo en una de las principales fuentes de energía. Además, menciona que los costos de O&M de los aerogeneradores se ven afectados negativamente por la ocurrencia de fallas. La detección temprana de estas fallas ayuda a reducir significativamente los costos. Como las fallas son difíciles de evitar incluso las graves, resultan en costosas paradas no programadas; por lo tanto, el objetivo de este proyecto es detectar cualquier fallo potencial lo antes posible.

4.1.1 Central Eólica Villonaco

El desarrollo de la energía Eólica en todo el mundo, ha comenzado en 1970 y en Ecuador continental inició en el año 2012, con la construcción de la CEV. A lo largo del tiempo ha ido disminuyendo los costos de generación y ha ido creciendo el rendimiento de Centrales Eólicas (Ayala et al., 2017).

La CEV está ubicada en el cerro del mismo nombre, en la provincia de Loja al sur de Ecuador, sus coordenadas son UTM 693030 E y 9556476 N. La CEV es el primer parque eólico en el Ecuador continental, con 16,5 MW de potencia nominal, situándose en relieve montañoso a aproximadamente 2 720 m.s.n.m (Ayala et al., 2017).

En la **Figura 1** se observa la CEV, que se constituye de 11 aerogeneradores de la marca GOLDWIND GW70, clase “S”, con tecnología *Direct Drive* y generador síncrono de imanes permanentes de 1,5 MW de capacidad nominal. El lugar se ubica en NNW SSE, en la cumbre del cerro Villonaco. La CEV se construyó con un costo aproximado de USD 41,8 millones, y opera desde 2013 (Hernández et al., 2016, citado en Maldonado-Correa et al., 2021).



Figura 1. Central Eólica Villonaco.

Fuente: Adaptado de La Central Eólica Villonaco [Fotografía], por Enrique (2021). Todos los derechos reservados [2022]. Tomado con permiso del autor.

La alta velocidad del viento la hace tener condiciones especiales, pues existe una velocidad media anual de viento superior a 10 m/s, baja densidad del aire de $0,89 \text{ kg/m}^3$ e intensidad de turbulencia de 0,15 (Reyes et al., 2017, citado en Maldonado-Correa et al., 2021).

4.1.2 Aerogenerador

Es una máquina que convierte la energía cinética del viento en energía eléctrica. Es decir, la fuerza del viento golpea las palas haciéndolas girar, convirtiéndola en energía mecánica, luego, se transfiriere al eje para llegar al generador, donde finalmente, se transforma en energía eléctrica.

4.1.3 Componentes de un aerogenerador

Los aerogeneradores tienen diferentes componentes y subsistemas, como se observa en la **Figura 2**, los componentes principales de un aerogenerador GW70/1500 son rotor,

generador, controlador, sistema de dirección de palas, sistema de orientación de la góndola, entre otros elementos.

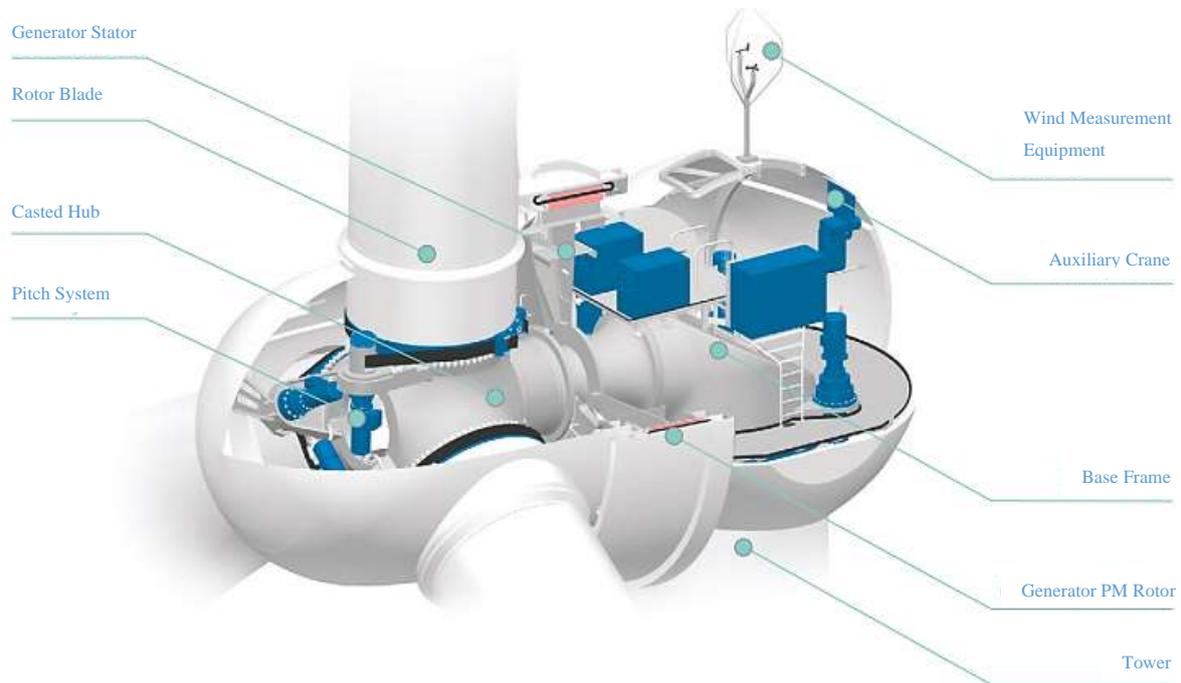


Figura 2. Componentes de un aerogenerador GW70/1500.

Fuente: Adaptado de Xinjiang Goldwind Science & Technology Co., Ltd (2011).

Seguido se hará una síntesis de los componentes de un aerogenerador.

4.1.3.1 Torre

Es fabricado de acero tubular, celosía de acero o hormigón. Este tiene la función de soportar la estructura de la turbina. Debido a que la velocidad del viento aumenta con la altura, se debe optar por torres altas para permitir la captura de energía por parte de las aletas de la turbina.

4.1.3.2 Aspas

Su forma es característica porque deben crear un diferencial de presión para hacer girar las aspas cuando el viento las atraviesa. La mayoría de aerogeneradores se constituyen de 3 palas, por razones como: en el caso de tener menor número de palas el rendimiento es bajo porque existe el fenómeno tambaleante y tensión en los componentes; sin embargo, en el caso de mayor número de palas aumentarían las cargas, el peso en la torre, la resistencia del viento y el costo.

4.1.3.3 Sistema de orientación de palas (*Pitch control*)

El rotor es la parte donde va conectada las palas al cubo; dentro de este se encuentra el sistema de Pitch, que controla el ángulo de ataque del viento con las palas para lograr el control de la velocidad del rotor.

4.1.3.4 Sistema de orientación de la góndola (*Yaw control*)

Este sistema se basa en el accionamiento de un motor, ajustando las ruedas y piñones para alinear el aerogenerador, según las condiciones del anemómetro y utilizando la información de velocidad del viento de forma adecuada.

4.1.3.5 Góndola

Es la carcasa que contiene todos los componentes para el funcionamiento del aerogenerador. Además, la góndola está diseñada para orientar la turbina eólica hacia la dirección del viento, lo que permite maximizar la captura de energía.

4.1.3.6 Veleta de viento

De acuerdo con la información proporcionada por la veleta de viento, el sistema Yaw ajusta la posición de la góndola.

4.1.3.7 Controlador

Este se encarga de arrancar y apagar el aerogenerador de forma controlada en función de la velocidad del viento.

4.1.3.8 Sistema eléctrico

Lo componen el transformador, fusibles, interruptores, cables, y demás conexiones.

4.1.3.9 Frenos

Un freno de disco mecánico sirve para detener el rotor, este solo se emplea en caso de falla del freno aerodinámico o para realizar mantenimiento.

4.1.4 Sistema SCADA

Este sistema se puede utilizar tanto para monitorear como para controlar sistemas industriales de forma remota y brindan una forma eficiente para recopilar y analizar datos en tiempo real. Cientos de miles de sensores pueden ser empleados en sistemas SCADA más grandes, generando grandes cantidades de datos.

En un estudio Maldonado-Correa et al. (2020), afirman que el sistema SCADA en aerogeneradores recolecta datos en intervalos de 1-10 minutos (frecuentemente cada 10 minutos), generando históricos de datos. Este sistema puede guardar más de 200 variables que incluyen velocidad del viento, velocidad del rotor, temperatura de la góndola, potencia activa, alarmas, ángulo de inclinación, temperatura del devanado o cojinete del generador, temperatura del transformador, temperatura de la caja de cambios, guiñada, temperatura ambiente, dirección del viento y otras señales eléctricas.

Los autores Blanco-M et al. (2018) hacen mención que las variables analógicas son las lecturas continuas de distintos sensores de los aerogeneradores en un determinado tiempo, el sistema SCADA proporciona variables discretas cuando existe una falla. Por ello, afirma que

los datos del SCADA son una fuente rica en información; por lo tanto, utilizando la monitorización automática se puede desarrollar sistemas y herramientas que ayuden a la toma de decisiones contribuyendo a la O&M.

El uso de datos SCADA tiene una gran ventaja debido a que se evita gastos innecesarios de instalación, mantenimiento de sensores y cables, dedicados a la adquisición de datos.

4.1.5 Datos de series temporales

Es una colección ordenada de mediciones tomadas cada cierto intervalo de tiempo, en un periodo a tratar.

4.1.6 Mantenimiento en aerogeneradores

Actualmente, los mantenimientos aplicados a los aerogeneradores son del tipo periódico clásico, basados en acciones y períodos determinados por el fabricante. Por consiguiente, se debe aplicar un plan de mantenimiento que tenga en cuenta de forma continua y durante toda la vida del aerogenerador, en sus condiciones de trabajo real, así como el efecto en componentes relevantes. Este tipo de acción se denomina mantenimiento basado en la condición que contribuirá a la fiabilidad, disponibilidad, mantenibilidad y seguridad de componentes del aerogenerador.

La nueva tendencia en el mantenimiento es la utilización de sistemas expertos y la inteligencia artificial.

4.1.7 Tipos de mantenimiento

4.1.7.1 Mantenimiento correctivo

Es un conjunto de actividades donde se repara o sustituye algún componente defectuoso y solo se lo realiza cuando ocurre un fallo. Su principal inconveniente es que el fallo puede ocurrir en cualquier momento, muchas veces el menos oportuno, debido a esfuerzos en el aerogenerador. Por lo tanto, cuando no se detectan estos fallos a tiempo, pueden influir en el daño de otros componentes. Otra dificultad radica en que, para realizar este mantenimiento, se debe tener repuestos en almacén, ya que, de no existir uno en el momento adecuado, se tendría que suspender la producción hasta que el proveedor envié el repuesto.

4.1.7.2 Mantenimiento preventivo

Son actividades programadas previamente, como inspecciones, pruebas, entre otras, están dirigidas a reducir la frecuencia y el impacto de los fallos. Sin embargo, tienen algunos inconvenientes, como el hecho de que se cometen cambios innecesarios en los componentes sin aprovechar su potencial al máximo, problemas iniciales de operación, una "mortalidad infantil" cuando se coloca una nueva pieza, una mayor necesidad de mano de obra intensiva y especializada por periodos de tiempo cortos.

4.1.7.3 Mantenimiento predictivo

Es el seguimiento y diagnóstico de un sistema que permite una intervención correctora inmediata como consecuencia de la detección de algún síntoma de fallo. Este proceso se basa en la condición de que un fallo se produce lentamente; es decir, se trata de seguir la evolución de futuros fallos.

4.1.8 Fallas en aerogeneradores

Según investigaciones realizadas por Dameshghi y Refan (2019), el acceso a los aerogeneradores y su mantenimiento de manera física es muy complejo y costoso. Por esta razón, el monitoreo basado en condiciones (CBM *abreviatura del inglés*) es muy apropiado para estas máquinas. Además, Dameshghi y Refan mencionan que el 75% de las fallas causan el 5% del tiempo de inactividad, mientras que el restante del 25% genera el 95% de tiempo sin operar, observando que los mayores índices de fallas se registran en subsistemas eléctricos y los mayores índices de inactividad en la caja de cambios (en caso de que el aerogenerador los tenga).

Según Blanco-M et al. (2018), identificar la salud en los aerogeneradores se vuelve crítico, ya que se reduciría un 25-35% de los costos de generación. Además, mencionan que la tarea requiere de mucho tiempo de un especialista porque se debe trabajar en cada aerogenerador individualmente, esto es el mantenimiento preventivo realizado cada 2500 a 5000 horas. Por consiguiente, Blanco-M et al. describen que la frecuencia de las operaciones del mantenimiento preventivo es insuficiente para predecir y detectar el estado del dispositivo, así anticiparse a las fallas que pueden ocurrir de manera inesperada causando costos significativos, debido a que generalmente las centrales eólicas suelen estar alejadas de áreas urbanas, y puede requerir varios días hasta la llegada del nuevo componente. Por esta razón, un análisis exhaustivo de un conjunto de aerogeneradores de una determinada central eólica no puede ser determinado por un humano.

Maldonado-Correa et al. (2020), mencionan que los componentes o subsistemas de los aerogeneradores a los que se pueden aplicar técnicas de Monitoreo basado en la Condición (CBM) son: caja de cambios, componentes eléctricos y electrónicos, palas y ángulo de inclinación, torre, tren de transmisión y cojinete, sistema de guiñada y sistema hidráulico siendo los más utilizados para CBM en la investigación que realizó.

Un enfoque presentado por Alejandro Blanco et al. (2019), con respecto al proceso clave para reducir la inactividad de los aerogeneradores es el uso de metodologías y herramientas que ayuden a estos procedimientos pueden ser muy beneficioso en centrales eólicas, con el fin de aumentar la producción y reducir los costos. Por ello, el CBM se basa en

la vigilancia continua, la detección de fallas emergentes y determinar tareas de mantenimiento antes de que se produzca la falla.

El estudio que realizó Artigao et al. (2018) sobre 13 aerogeneradores tanto de tierra como de mar, reveló que los conjuntos más críticos son el sistema de control, la caja de cambios (caso de tener), el sistema eléctrico, el generador y las palas, ya que representan un mayor porcentaje de fallas y tiempo de mantenimiento. Por otro lado, mencionan que los conjuntos más confiables son la estructura y los sensores. Además, Artigao et al. afirman que el principal componente en fallar con mayor frecuencia es el sistema eléctrico.

Las fallas frecuentes en varios aerogeneradores en tierra (*onshore*), presentadas por Hossain et al. (2018), como se puede observar en la **Figura 3**, se encuentran: el convertidor, el sistema de control, el rotor, y los sensores, los cuales son responsables del 63% de las fallas en los aerogeneradores.

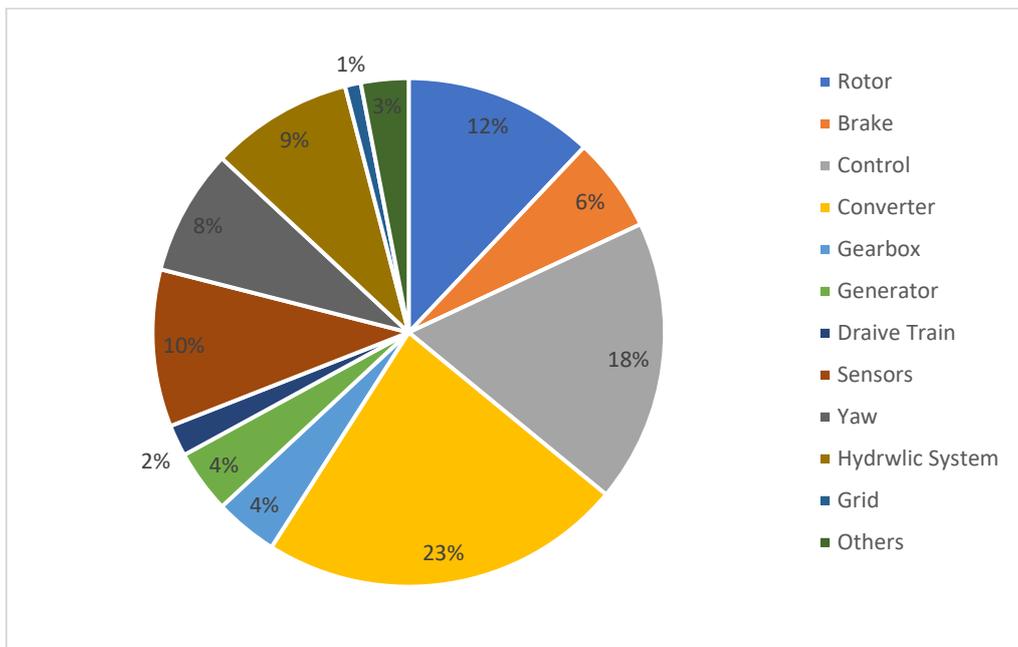


Figura 3. Porcentaje de fallas en los componentes de los aerogeneradores onshore.

Fuente: Adaptado de Hossain et al. (2018)

4.1.8.1 Convertidor electrónico

Hossain et al. (2018), mencionan que la confiabilidad de este elemento se vuelve más compleja con el aumento de la potencia nominal de los aerogeneradores, siendo la temperatura, la vibración y la humedad, los factores principales que causan las fallas en los convertidores de potencia. Además, argumentan que existen 3 componentes donde se presentan más fallas como: condensadores, circuito de placa impreso (PCB *abreviatura del inglés*) y transistor bipolar de puerta aislada (IGBT *abreviatura del inglés*). La **Figura 4** muestra el porcentaje de fallas que ocurren en estos componentes.

Desde el punto de vista de Hossain et al. (2018), las fallas en el capacitor incluyen: apertura/cortocircuito, materiales de electrodos que migran por el dieléctrico y forman caminos conductores, dieléctrico en descomposición y aumento del factor de disipación. Asimismo, aseguran que las fallas del PCB son: líneas de metal mal enterradas rotas, grietas de trazas, corrosión, desalineación de componentes, delaminación de tableros y soldadura en frío; y las fallas en los módulos IGBT, incluyen: despegue del cable de unión de cortocircuito, fallo en el encendido de la puerta por grietas y fatiga en la soldadura. Hossain et al. también explican que para monitorear la degradación del IGBT se usan variables cómo voltaje de saturación de colector-emisor, voltaje de umbral del emisor de puerta, resistencia activa y la resistencia interna.

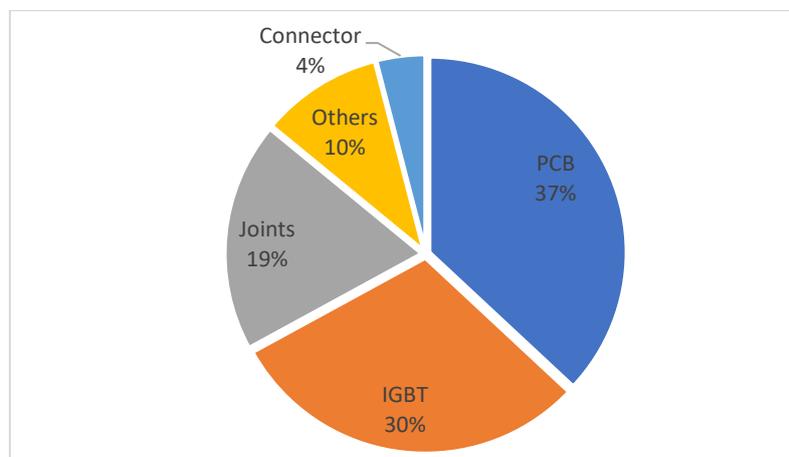


Figura 4. Fallas en los componentes del convertidor electrónico de potencia.
Fuente: Adaptado de Hossain et al. (2018)

4.1.8.2 Generador eléctrico

El generador puede tener 2 tipos de fallas, eléctricas y mecánicas. Entre las fallas eléctricas comprendidas están el circuito abierto por aislamiento del estator o rotor y desequilibrio eléctrico. Por otra parte, entre las fallas mecánicas se incluyen el eje roto, falla en rodamiento y desequilibrio de la masa del rotor. De la literatura revisada por Yucai et al. (2016, citado por Hossain et al., 2018), la falla más común es el cortocircuito de las bobinas de un generador, por lo que, el sensor de temperatura es el que puede detectar anomalías en el devanado. Además, Hossain et al. (2018) hace referencia que el desequilibrio eléctrico del estator se puede detectar mediante el cambio en el contenido armónico de las señales eléctricas.

4.1.8.3 Sistema hidráulico

Este sistema se acciona mediante motores, aumentando la presión de fluido hidráulico para ajustar el ángulo de paso de la pala, además de ajustar la posición de la góndola y controlar el freno mecánico. Entre las fallas frecuentes, se encuentran la fuga de aceite y fallas de bloqueo

de válvulas deslizantes (Hossain et al., (2018). Las señales de los sensores de presión y nivel se utilizan para diagnosticar estas fallas.

4.1.8.4 Sensores

Hossain et al. (2018), en su investigación recopiló que las fallas más frecuentes en los sensores se presentan en los transductores de temperatura, voltaje, corriente y par, ocurriendo el 14% de fallas en estos elementos. Estas pueden ocurrir por mal funcionamiento de software causando degradación en el rendimiento de la turbina, debido a que, estos elementos envían información al sistema de control de subsistemas mecánicos y eléctricos, lo que puede provocar una disminución en la eficiencia y hasta el apagado de la turbina eólica.

4.1.8.5 Sistema de control

Este sistema es muy importante para regular las operaciones en un aerogenerador. Puede presentar fallas en el hardware como en el software. Las fallas en el hardware incluyen fallas en sensores, actuadores, placa de control y enlaces de comunicación (Hossain et al., 2018). Las fallas en el software incluyen desbordamiento de buffer (buffer overflow), fugas de recursos y falta de memoria.

4.1.8.6 Sistema de orientación de palas.

El sistema de control de paso es muy importante para la recolección de energía, la pérdida de viento y frenado aerodinámico. Por lo tanto, una falla en este sistema puede tener consecuencias devastadoras cuando los vientos fuertes amenazan la capacidad del aerogenerador para operar de manera segura (Olabi et al., 2021). Generalmente, el sistema es operado por un motor eléctrico o actuador hidráulico; sin embargo, los motores eléctricos proporcionan tiempos de reacción veloces, por lo que tienen mayor rigidez y confiabilidad en comparación con los sistemas hidráulicos; en cambio, el sistema hidráulico da como resultado inestabilidad operativa y la activación temprana del freno puede debilitar el modo de funcionamiento deseado.

4.2 Capítulo II: Procesamiento de datos

4.2.1 *Preprocesamiento de datos*

En muchos conjuntos de datos se presentan inconsistencias o ausencias, por lo que, si no se atiende adecuadamente puede ocasionar errores en los resultados, de ahí su importancia (Pertuz, 2022).

Cuando el CBM se realiza con base en datos del sistema SCADA, la resolución de los datos no es la adecuada, para el procesamiento de señales directamente en tiempo y frecuencia. En general, hay pocas estrategias para usar datos SCADA, siendo una de ellas la búsqueda de una señal de tendencia en las variables (Dameshghi & Refan, 2019).

En el estudio que hizo Hossain et al. (2018), da a conocer las variables y componentes comprometidos para predecir fallas; esto se resume en la **Tabla 1**, que incluye los componentes, fallas, señales de monitoreo y métodos de CM utilizados generalmente.

Tabla 1. Resumen de fallas, señales y métodos utilizados para procesar datos.

Componente	Fallas ocurridas	Señales monitorizadas	Métodos de procesamiento
Rotor	1. Fatiga	1. Vibración	1. Transformada de Hilbert
	2. Grieta	2. AE	2. Muestreo síncrono
	3. Rugosidad (superficie)	3. Presión	3. FFT
	4. Asimetrías	4. Torque	4. Transformada de Wavelet
	5. Rigidez reducida	5. Eléctrico	5. Método basado en modelos
	6. Deformación	6. NDT	6. AI
Rodamientos	1. Rugosidad (superficie)	1. Vibración	1. Transformada de Hilbert
	2. Fatiga, Grieta, Rotura	2. AE	2. Muestreo síncrono
	3. Carrera exterior/interior	3. Eléctrico	3. FFT
	4. Bola o Jaula	4. Temperatura 5. Aceite	4. Métodos basados en modelos 5. AI
Eje principal	1. Corrosión	1. Vibración	1. Muestreo síncrono
	2. Grieta	2. Torque	2. Transformada de Hilbert
	3. Desalineación, falla de acoplamiento	3. Eléctrico	3. FFT
Sistema hidráulico	1. Fuga de aceite	1. Nivel de presión	1. Comparación de umbral
	2. Obstrucción de la válvula deslizante		
Freno mecánico	1. Desgaste del disco/pinza	1. Vibración	1. Estadística
	2. Grieta en disco	2. Temperatura	2. FFT
	3. Falla de sección hidráulica	3. Eléctrico	3. Métodos basados en modelos
	4. Fallo de motor		4. Comparación de umbral
Torre	1. Corrosión	1. Vibración	1. FFT
	2. Grieta		2. Métodos basados en modelos
	3. Daño estructural		
Generador	1. Circuito abierto/cerrado	1. Vibración	1. Muestreo síncrono
	2. Daño de aislamiento	2. Torque	2. Transformada de Hilbert
	3. Desbalance	3. Temperatura	3. Envolvente
	4. Rotura de barra de rotor	4. Parámetros de aceite	4. Estadística
	5. Eje doblado	5. Eléctrico	5. FFT
	6. Falla de rodamiento		6. Transformada de Wavelet
	7. Excentricidad del entrehierro		7. Métodos basados en modelos
	8. Fallas de imanes Masa del rotor		8. Comparación de umbral AI
	9. desbalanceada		
Convertidor de potencia	1. Capacitor	1. Temperatura	1. Estadística
	2. PCB	2. Eléctrico	2. Métodos basados en modelos
	3. Semiconductor		3. Comparación de umbral 4. AI.
Sensores	1. Sensor	1. Todas las señales relacionadas	1. Estadística
	2. Hardware de procesamiento de datos		2. Métodos basados en modelo
	3. Comunicación		3. AI
	4. Mal funcionamiento del software		
Sistema de control	1. Sensor	1. Todas las señales relacionadas	1. Estadística
	2. Actuador		2. Métodos basados en modelo
	3. Controlador		3. AI
	4. Comunicación		
	5. Malfuncionamiento de Software		

Fuente: Adaptado de Hossain et al. (2018).

4.2.2 Integridad de los datos

Se refiere respecto a la limpieza de valores atípicos, valores nulos y bajas características de varianza. Es un paso importante donde los datos deben ser corregidos o eliminados antes de planificar y construir los modelos. Para llevar a cabo esta tarea, se utilizará cuadernos de *Google Colab* y el *IDE Spyder* en el lenguaje de programación Python, así como las librerías de visualización y exploración de datos como: Pandas, Matplotlib y Seaborn.

La fase de limpieza de datos para identificar fallas, según Santolamazza et al. (2021), se debe realizar siguiendo los siguientes pasos:

- Eliminación de muestras con valores faltantes.
- Eliminación de muestras en las que la potencia de salida del aerogenerador es cero.
- Retiro de muestras con valores de las variables que estén fuera del rango normal.

4.2.2.1 Datos faltantes

Para Pertuz (2022), es común que al recopilar datos se encuentren que falte alguno de ellos; por ello, se debe prestar atención, ya que esto puede provocar problemas en el algoritmo utilizado. Además, Ozdemir y Susarla (2018) consideran que completar los datos donde antes faltaban puede ir desde eliminar una columna por completo hasta soluciones muy complejas.

4.2.2.1.1 Eliminación de filas dañinas

Es lo más simple que se puede realizar para manejar los datos faltantes, al hacerlo solo quedarán los datos de puntos completos (Ozdemir y Susarla, 2018).

4.2.2.1.2 Relleno de valores perdidos

Es un método complejo, es el acto de completar los valores faltantes con cantidades numéricas que de alguna manera se determinan a partir de los datos existentes; por ejemplo, con el valor promedio de la columna (Ozdemir & Susarla, 2018).

4.2.3 Ingeniería de características

La ingeniería de características implica la comprensión y transformación de variables en relación al aprendizaje supervisado y no supervisado (Ozdemir & Susarla, 2018).

4.2.3.1 Selección de características

Los procedimientos para un mejor rendimiento de los modelos de predicción de fallas, requieren la selección de variables importantes, debido a la gran cantidad de sensores que existen (Marti-Puig et al., 2019). Se debe tomar en cuenta que las variables de entrenamiento y de prueba son las mismas (Vilema Lara, 2022).

Para Ozdemir y Susarla (2018), la extracción inteligente de variables se puede mejorar el comportamiento del modelo al eliminar datos redundantes, ruidosos e irrelevantes, así como reducir el tiempo de entrenamiento y predicción.

4.2.3.1.1 Método de eliminación recursivo de características (RFE por sus siglas de Recursive Feature Elimination)

Este método elige la entrada que más se ajusta al modelo, con el objetivo de eliminar atributos que no son dependientes entre sí, y que tienen correlación lineal entre ellos. De acuerdo con Rashid & Batunlu (2021), es un método determinista que consiste en ejecutar el modelo varias veces e ir seleccionando diversas variables; es decir, en función de los resultados se elige las mejores variables.

4.2.3.1.2 Método estadístico de correlación

El coeficiente de correlación de Pearson mide la relación lineal entre columnas. Este valor de coeficiente varía entre -1 y 1, donde estos valores representan una correlación fuerte entre columnas (Ozdemir & Susarla, 2018). En este método se puede digerir si cualquier característica esta más relacionada con la respuesta más útil será; además, se lo usa para determinar interacciones y redundancias en variables. Esto puede ayudar a reducir el sobreajuste del modelo. El autor Mendapara (2021) define que, matemáticamente, es la covarianza entre dos señales que se divide por la multiplicación de la desviación estándar y está definida cómo (Véase **ecuación (1)**):

$$\rho(X, Y) = \frac{COV(X, Y)}{\sigma_X \sigma_Y} \quad (1)$$

4.2.3.1.3 Método estadístico de prueba de hipótesis

Esta metodología aplica pruebas estadísticas a cada una de las variables por separado seleccionando las mejores en un conjunto de datos (Ozdemir & Susarla, 2018). Es un método parecido al de correlación, pero en este caso se usa un valor-p (un decimal entre 0 y 1), pero en este caso mientras menor sea el valor-p, mayores posibilidades de que la característica tenga relevancia fuerte.

4.2.3.1.4 Métricas de selección basado en árboles

Este método utiliza arboles de decisión que comienzan desde el nodo raíz y elige con fuerza la división optima, en cada cruce que mejora una cierta métrica de pureza del nodo, al hacerlo dichas métricas tienen una noción de la importancia de las variables (Ozdemir & Susarla, 2018).

4.2.3.1.5 Modelos lineales y regularización

Los modelos lineales como regresión lineal, regresión logística, las máquinas de soporte, y otros colocan coeficientes al frente de las variables representando la misma pendiente (Ozdemir & Susarla, 2018). Para ello, se debe conocer acerca de la regularización, que se define como el método de imponer restricciones adicionales para evitar el sobreajuste, ayudando a generalizar los datos y evitar que las variables estén correlacionadas entre sí.

Ozdemir y Susarla (2018) proponen que, si las funciones son cuantitativas, los métodos adecuados para la selección de variables son: confiar en las correlaciones y selectores basados en modelos; además, mencionan que, si se trata de un problema de clasificación binaria, es muy probable el uso de clasificación de vectores de soporte, y un poco de exploración de datos ayudaría a la selección manual de funciones, tomando en cuenta el dominio que se tiene en el campo.

4.2.3.2 Escalado y normalización de características

Es una tarea importante para que todas las variables de un conjunto de datos estén bajo la misma escala, generalmente en un rango entre 0 y 1, esto se debe hacer tanto para el conjunto de prueba y entrenamiento a fin de no provocar un sesgo aleatorio (Pertuz, 2022). Por otra parte, Thoresen (2019) considera que el escalado es parte del preprocesamiento en una distribución normal; pero debe hacer con cuidado, en variables muy sesgadas y que tienen valores extremos, porque producirá resultados deficientes. También, Rashid y Batunlu (2021) destacan que los datos a diferentes escalas son muy difíciles de visualizar y puede degradar el rendimiento de un modelo de aprendizaje automático, aumentando el tiempo de la predicción. La normalización se usa frecuentemente definida por la **ecuación (2)**:

$$x_{normalizada} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2)$$

4.2.4 Datos de entrenamiento, validación y prueba

Según el proyecto de investigación realizada por Eriksson (2020), es común dividir los datos en conjuntos de entrenamiento, validación y prueba. Los datos de entrenamiento son usados para ajustar el modelo inicial, mientras que los de validación para proporcionar métricas imparciales sobre la capacidad de generalización el modelo con los parámetros de elección. Eriksson afirma que cuando se entrena y se encuentra el mejor modelo, los datos de prueba son empleados para evaluar que tan bien se generaliza el modelo. Es frecuente dividir los datos en dos partes, la primera parte para entrenamiento y validación, y la otra para datos de prueba, como regla general con proporciones de 75%-25%, 70%-30% o 80%-20%, siempre y cuando se tenga en cuenta la cantidad de datos disponibles (Pertuz, 2022).

4.3 Capítulo III: Estado del arte

4.3.1 *Inteligencia artificial*

De acuerdo con Rouhiainen (2018), la Inteligencia Artificial (IA) es la capacidad de las máquinas de usar algoritmos, aprender de los datos y usarlo en la toma de decisiones, tal como lo haría un humano. No obstante, estas máquinas pueden analizar grandes volúmenes de datos y pueden cometer menos errores que los humanos.

Javaid et al. (2022), en su libro destacan que la IA contribuye al desarrollo de la industria, enfocándose en mejorar los productos, aumentar la productividad y reducir los costos de operación. Además, afirman que la IA proporciona información adecuada para la toma de decisiones y alertas de fallos en el sistema.

Según Rouhiainen (2018), la IA se puede aplicar en diversas situaciones cómo:

- Reconocimiento y clasificación de imágenes.
- Mantenimiento predictivo.
- Protección contra amenazas en cibernética.
- Detección y clasificación de objetos.
- Mejoras de estrategias comerciales.
- Procesamiento eficiente y escalable de datos.

4.3.2 *Aprendizaje automático (Machine Learning)*

Es un conjunto de técnicas que obtienen información y relaciones de comportamiento de datos utilizando métodos matemáticos y estadísticos (Ozbas et al., 2019). El autor Sandoval Serrano (2018), define el aprendizaje automático como una rama de la IA, que se encarga de generar algoritmos que tienen la capacidad de aprender, donde se tendrá que alimentar al algoritmo con una enorme cantidad de datos para que este relacione los datos y sepa que hacer.

Actualmente, existen dos métodos de aprendizaje automático, el supervisado y el no supervisado. El aprendizaje automático supervisado es con mucho el más común en una amplia gama de casos de uso en la industria (Vives et al., 2020).

En la **Figura 5** se muestra el proceso general, que consiste, primeramente, en obtener datos originales desde una fuente confiable, luego limpiarlos y normalizarlos. Después de diseñar el modelo, este analiza los datos de entrenamiento, aprendiendo de su relación, para posteriormente resolver el problema con los datos de prueba. Finalmente, se comparan los datos reales con los de prueba, determinando la precisión del modelo (Forootan et al., 2022).



Figura 5. Modelo general de aprendizaje en un modelo de aprendizaje automático.
Fuente: Adaptado de Forootan et al. (2022).

4.3.2.1 Detección de fallas en aerogeneradores con aprendizaje automático

El aprendizaje automático en la detección de fallos en aerogeneradores se centra en la detección de anomalías como en la clasificación de éstos. Vives et al. (2020), argumentan que esta técnica permite detectar el fallo de manera temprana, incrementando el nivel de fiabilidad y seguridad del sistema, al tiempo que se facilita la toma medidas para el mantenimiento predictivo.

Ding (2019) argumenta que detectar una anomalía es el primer paso para informar acciones adecuadas, debido a que a menudo una anomalía simplemente revela el síntoma del problema. Sin embargo, múltiples causas profundas pueden conducir a un mismo síntoma por lo que es inevitable un seguimiento de diagnóstico.

En el trabajo realizado por Chen et al. (2018), introdujeron un modelo generalizado para la detección de condiciones defectuosas en los aerogeneradores, utilizando predicción combinada y entropías de la información; además, destacan que el modelo de predicción de combinación lineal es efectivo para parámetros SCADA, presentando mejor la precisión y rendimiento de generalización. Así mismo, en Liu et al. (2020) proponen análisis y determinación de indicadores de fallas en el convertidor, el generador y el sistema de pitch, donde aplicaron una red neuronal convulsional con ResNet50 y un clasificador de máquinas de vectores de soporte y un método gráfico de radar propuestos. En la investigación que realizó Carbajo y Bosch (2022), se presentaron las capacidades de aprendizaje automático para la predicción de fallas en turbinas eólicas, así como una técnica para predecir diferentes tipos de fallas donde comparó el rendimiento de clasificadores K-Nearest Neighbour y Random Forest.

4.3.3 *Aprendizaje profundo (Deep Learning)*

Este utiliza metodología empleada en el aprendizaje de automático, mediante el empleo de redes neuronales que se componen de un número de niveles jerárquicos caracterizados por constitución de múltiples capas de procesamiento no lineal. Vives et al. (2020), en su investigación encontraron que muchos algoritmos de aprendizaje profundo, han sido diseñados para abordar problemas no supervisados, mediante un entrenamiento previo sin supervisión distribuido por capas seguido de un ajuste supervisado, debido a la dificultad que presenta la alta dimensionalidad de las variables aleatorias que se modelan en el aprendizaje no supervisado. En la actualidad, existe varias arquitecturas de aprendizaje profundo tales como las redes neuronales, representaciones gráficas o mecanismos de atención (Esta última área se encuentra en constante evolución).

Chatterjee y Dethlefs (2021) en su revisión sobre la IA para O&M de aerogeneradores describió las principales áreas de enfoques probables en la industria eólica en función de las tendencias, éxitos y desafíos actualmente, donde afirman que el aprendizaje profundo recibiría un enfoque mayoritario para el año 2022.

4.3.4 *Tipos de aprendizaje de aprendizaje profundo*

Existen 3 maneras de entrenar modelos de aprendizaje profundo, dependiendo si los datos están etiquetados o no:

- Aprendizaje supervisado
- Aprendizaje no supervisado
- Aprendizaje por refuerzo

4.3.4.1 Aprendizaje supervisado

En este aprendizaje, los datos se encuentran en forma de ejemplos con etiquetas, lo que permite al algoritmo predecir al aproximarse a la exactitud de la relación entre los ejemplos y las etiquetas. Al aprender la relación entre los datos, el algoritmo será capaz de observar un nuevo ejemplo y predecir a qué etiqueta pertenece.

Según Eriksson (2020), este método puede ser utilizado cuando cada muestra de conjunto de datos utilizado para el entrenamiento del modelo tiene un conjunto de valores de entrada x conocidos, al igual que los valores de salida. Este algoritmo se utiliza para encontrar una función de mapeo $f(x) = y$. Eriksson menciona que es importante comprender la diferencia entre un problema de clasificación y regresión. Un problema de clasificación es cuando se crea una función de mapeo tal que, mapee la entrada a un valor discreto, mientras

que un problema de regresión genera un mapeo de función a una variable continua. En consecuencia, el mantenimiento predictivo puede ser de clasificación o regresión.

4.3.4.2 Aprendizaje no supervisado

El aprendizaje no supervisado es opuesto al aprendizaje supervisado, en este los datos no presentan etiquetas. En este caso, el algoritmo deberá aprender a agrupar los datos de tal manera que pueda reconocer a que grupo pertenecen los datos nuevos ingresados.

Este tipo de aprendizaje se usa para realizar inferencias a partir de datos que no están etiquetados, debido a que se desconocen los datos de salida, no se puede aplicar una regresión directamente. Esta técnica agrupa valores de entrada para encontrar patrones desconocidos o estructuras subyacentes en los datos; además, incluye un análisis exploratorio y reducción de la dimensionalidad (Eriksson, 2020).

4.3.4.3 Aprendizaje por refuerzo

El aprendizaje por refuerzo consiste en básicamente aprender de los errores. Para ello, se necesita un agente y un entorno conectados a través de un circuito de retroalimentación. Para Janiesch et al. (2021), este tipo de aprendizaje, en lugar de proporcionar pares de entrada y salida, se describe el estado actual del sistema, donde se especifica un objetivo proporcionando una lista de acciones permitidas y restricciones. Por lo tanto, el modelo debe experimentar el proceso usando el principio de prueba y error, maximizando una recompensa si el resultado es mejor.

4.3.5 *Detección de anomalías con aprendizaje profundo*

Según L. Xu et al. (2022), en su publicación mencionan que el aprendizaje profundo, puede mejorar la información oculta del conjunto de datos, lo que tiene mejor efecto en la detección de anomalías, lo que atrae a investigadores a estudiar el tema. Por ejemplo, el método de aprendizaje profundo utilizado en la detección de anomalías es el AutoEncoder publicado por Pumsirirat y Yan (2018); sin embargo, este método no tiene una regularización fuerte, lo que facilita un sobreajuste (Overfitting); por ejemplo, si hay muchos datos anormales aprenderá patrones anómalos.

Chatterjee y Dethlefs (2021) realizaron un análisis cuantitativo del uso de la IA en la industria Eólica, y obtuvieron que la investigación está avanzando hacia el aprendizaje profundo especialmente la IA explicable, siendo el modelo de transformadores uno de ellos.

4.3.6 *Mecanismo de atención: El Modelo del Transformador (The Transformer Model)*

El modelo de transformador fue publicado por primera vez por Vaswani et al. (2017), y fue utilizado por primera vez en la industria Eólica por Meng y Xu (2019), que propone un codificador-decodificador de atención de múltiples cabezas con ajuste estacional y de

tendencia. Este modelo resulta ser mejor que modelos como el bosque aleatorio y árbol de decisión de aumento de gradiente. Niu et al. (2021), publicó que el mecanismo de atención se ha convertido en uno de los conceptos más importantes en el aprendizaje profundo, debido a que mejora en gran medida la eficiencia y precisión del procesamiento de información perceptual.

Los modelos de transducción de secuencia dominante se basan en complejas redes neuronales convolucionales o recurrentes que incluyen un codificador y un decodificador unidos a través de un mecanismo de atención. Vaswani et al. (2017), presentaron el modelo del transformador, una nueva arquitectura de red simple basada únicamente en mecanismos de atención, prescindiendo de redes de recurrencia y las convolucionales. Esta arquitectura fue principalmente desarrollada para la traducción automática, demostrando ser superior en calidad, ya que es paralelizable y requiere menos tiempo para entrenarse.

El modelo de transformador se ha aplicado con éxito al Procesamiento de Lenguaje Natural (PNL *abreviatura del inglés*) y demuestra sus capacidades para el modelado de datos de series temporales. Se han publicado algunos trabajos sobre el modelo de transformadores como Shaw et al. (2018), que propusieron el concepto de posición relativa para que el transformador pueda adaptarse a secuencias de diferentes longitudes. Dai et al. (2019), introdujeron un mecanismo de recurrencia a nivel de segmento, para establecer una conexión entre cada segmento de texto, permitiendo el modelo poder capturar más dependencias lejanas.

El modelo de transformador de Vaswani et al. (2017), puede capturar de manera eficiente la correlación a largo plazo dentro de la secuencia y realizar la generación de entrada a salida de extremo a extremo. Este modelo se forma de una pila (stack) de codificadores y decodificadores basados en autoatención.

4.3.6.1 Codificador

Se compone de seis capas idénticas, cada una de ellas con dos subcapas. La primera es un mecanismo de atención de múltiples cabezas y la segunda es una posición simple de red prealimentada (feed-forward) conectado a la posición. Se emplea una conexión residual alrededor de cada una de las dos subcapas, seguido de la normalización de capas. Esto se implementa mediante la salida de cada subcapa como $LayerNorm(x + Sublayer(x))$; donde $Sublayer(x)$ es la función implementada por la subcapa. Par facilitar el modelo y las conexiones residuales, todas las capas del modelo, así como su incrustación, producen una salida de dimensión $d_{model} = 512$ (Vaswani et al., 2017).

4.3.6.2 Decodificador

Se compone de 6 capas idénticas. Cada una de ellas tiene dos subcapas de cada decodificador y se inserta una tercera subcapa con múltiples cabezas de atención sobre cada salida del decodificador. Además, se emplea conexiones residuales alrededor de cada una de las subcapas, seguido de una capa de normalización. El enmascaramiento combinado de incrustaciones de salidas se utiliza para evitar que las posiciones se atiendan en posiciones futuras. Esto se logra compensando en una posición que depende solo de las salidas conocidas en posiciones anteriores a i (Vaswani et al., 2017). En la **Figura 6**, se muestra la arquitectura del transformador diseñada para traducción automática por Vaswani et al. (2017).

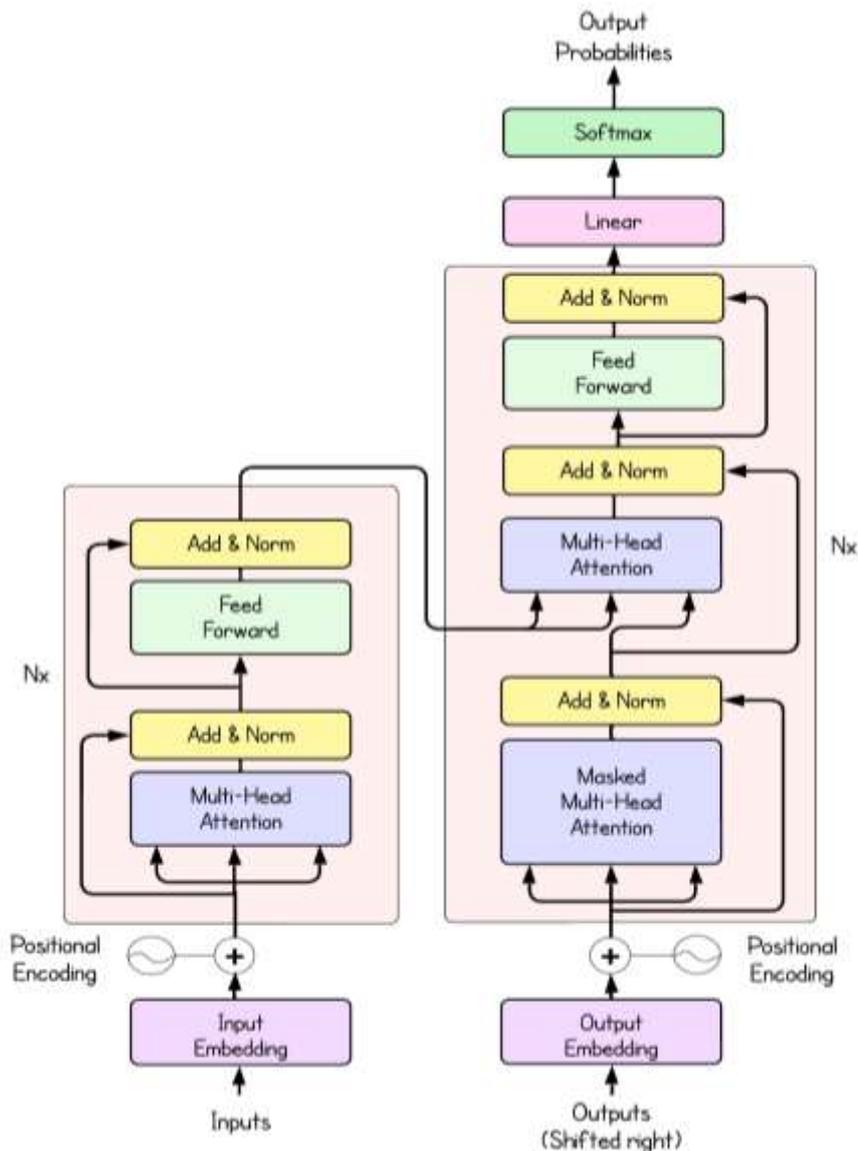


Figura 6. Arquitectura del transformador.
Fuente: Adaptado de Vaswani et al. (2017).

En resumen, según L. Xu et al. (2022), el codificador comprime la secuencia de entrada para generar codificación semántica y matriz de autoatención; el decodificador, al basarse en la autoatención, enfoca la atención en la información efectiva de la entrada de secuencia y genera el vector objetivo a través de la red residual y una red neuronal de avance. Además, L. Xu et al. argumenta que la entrada al transformador es una matriz multidimensional, que puede ser una serie de tiempo multivariante bajo una ventana deslizante (sliding), lo que permite al mecanismo de autoatención analizar el contexto de secuencia. Por otro lado, el marco de codificación-decodificación utiliza la salida del codificador para la entrada del decodificador realizando la conversión de modalidades (L. Xu et al., 2022).

4.3.6.3 Atención

Según Vaswani et al. (2017), una función de atención se puede caracterizar como la asignación de consulta (*Query: Q*), clave (*Key: K*) y Valor (*Values: V*), los cuales tienen como salida una suma ponderada de valores, siendo el peso asignado a cada valor calculado por una función de compatibilidad.

El cálculo de atención para todos los tokens se puede expresar como un cálculo de matrices, utilizando la función de *softmax*, que es útil para el entrenamiento debido a la matriz computacional de optimización de operaciones que se calculan rápidamente de forma matricial. Las matrices *Q*, *K* y *V*, se definen como las matrices donde *i*-énimas filas son vectores q_i , k_i , y v_i respectivamente. Según Vaswani et al. (2017), estas tres matrices de variables se obtienen proyectando los datos de entrada mediante tres cálculos matriciales, dado como la **ecuación (3)**.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{(d_k)^{\frac{1}{2}}}\right) \quad (3)$$

Las matrices *Q* y *K* son los operadores que se emplean para calcular la similitud del elemento actual en la secuencia con los otros elementos. La matriz *V* contiene información contenida en los propios elementos. Calculando las tres variables, cada elemento en la secuencia interactúa con elementos en otras posiciones, permitiendo aprovechar la información contextual por completo para generar la salida.

El funcionamiento en paralelo del mecanismo de autoatención hace que la distancia entre cualquier par de palabras en una secuencia sea como uno y captura información contextual de largo alcance. Este modelo puede resolver problemas de dependencia a largo plazo y exceso de consumo de capacidad de red, causado por la operación en serie donde cada

codificación oculta de un token, depende de la salida de la codificación oculta del token y de la salida de codificación del token anterior.

Hassid et al. (2022), consideran que el mecanismo de atención es la columna vertebral de la arquitectura Transformador, debido a que, contextualiza la entrada mediante el cálculo de matrices de atención específicas de la entrada, destacando así su poder y elegancia.

4.3.7 Detección de anomalías con la arquitectura de transformadores

Según L. Xu et al. (2022), la detección de anomalías tiene como objetivo detectar puntos de datos o fragmentos que no se ajustan a patrones de comportamiento esperados en datos que cambian rápidamente. En la industria, los patrones no conformes se denominan anomalías, valores atípicos u observaciones no acordes. La detección de anomalías se usa principalmente en áreas de la energía, las finanzas, la seguridad de la red, monitoreo de la salud, comercio, entre otros.

Dehghani et al. (2018), proponen un Transformador Universal, que introdujo la codificación de paso de tiempo, posición y tiempo, reemplazando la capa de avance con la función de transición, mejorando la versatilidad del transformador. Wu et al. (2020), presentaron el uso de un mecanismo de autoatención para aprender patrones y dinámicas complejas en datos de series temporales, tanto univariadas y multivariadas. S. Wu et al. (2020), dieron a conocer un modelo de detección de anomalías nuevo para series temporales, el Transformador Disperso Adversario (AST *abreviatura del inglés*), este utiliza el transformador como generador para aprender mapas de atención dispersa para predecir series temporales y un discriminador para mejorar el rendimiento de la predicción a nivel de secuencia.

Para Shang et al. (2021, citado por L. Xu et al., 2022), las anomalías se pueden dividir en tres tipos: puntuales, anomalías contextuales y anomalías colectivas. Debido a la heterogeneidad de la O&M presenta un desafío para mejorar la precisión de la detección mediante el análisis automático y resumir patrones anormales. Los métodos estadísticos tienen dificultad para identificar anomalías en series de tiempo, debido a que tienen una capacidad limitada para extraer información contextual. Por ello, los investigadores tratan de usar el aprendizaje automático para procesar gran cantidad de datos. Un método de aprendizaje no supervisado para anomalías proporciona una solución viable para deshacerse del costo manual. Karadayi et al. (2020, citado por L. Xu et al., 2022) mencionan que los modelos de detección de anomalías sin supervisión monitorean series de tiempo al predecir y reconstruir la serie temporal, y calcular la desviación entre el valor real y el predicho.

Una red de transformadores presentada por Marino et al. (2022), para detectar anomalías en tráfico de red, mejorando la interpretabilidad en la estructura de redes

informáticas, se entrena de manera autosupervisada y no requiere datos etiquetados. Este modelo fue evaluado para el Sistema de Control Industrial, identificando con éxito las anomalías.

También, según una investigación importante presentada por Tuli et al. (2022), se presenta un modelo de transformador denominado TranAD, que es una red de transformadores profundos para la detección de anomalías utilizando codificadores de secuencia basados en la atención, utilizando el meta-aprendizaje independiente del modelo (MAML *abreviatura del inglés*), que permite entrenar el modelo con datos limitados. Tuli et al. entrenaron el modelo en seis conjuntos de datos superando a métodos de última generación en rendimiento al detectar las anomalías, aumentando los puntajes de F1 hasta un 17%.

Un enfoque valioso lo presenta Xu et al. (2021), mediante un modelo de transformador de anomalías con un nuevo mecanismo de atención de anomalías, para calcular la discrepancia de asociación, gracias al diseño de una estrategia de minimax para amplificar la diferencia entre un punto anómalo y otro normal. Este modelo logra resultados de vanguardia en seis conjuntos de datos de series temporales no supervisadas aplicadas a monitoreo de servicios, exploración espacial y terrestre, y tratamiento de agua.

Un estudio reciente lo presentan L. Xu et al. (2022), donde muestran un modelo de arquitectura de transformadores basado en Redes Adversarias Generativas (GAN *abreviatura del inglés*), para detectar anomalías en datos de series de tiempo usado en la red eléctrica y mantenimiento de detección de flujo, diagnóstico de fallas y otras aplicaciones. A menudo, las series carecen de definiciones etiquetadas estrictas y los métodos existentes a menudo sufren la necesidad de aplicar normas rígidas debido a la incapacidad de manejar datos de alta dimensión y cálculos que consumen mucho tiempo. TGAN-AD como lo nombra L. Xu et al. a su modelo puede extraer características contextuales de series de tiempo, y su discriminador puede detectar datos anómalos. Las puntuaciones de anomalías se calculan mediante el generador y el discriminador ofreciendo rendimientos altos en F1 Score y alta eficiencia en la elección de los hiperparámetros.

Díaz (2020) aplicó la arquitectura de transformadores para el monitoreo predictivo de procesos comerciales y detección de anomalías, obteniendo resultados comparables a pesar de que el modelo no es una red tan compleja como LSTM o GRU.

En resumen, el modelo de transformador ha demostrado ventajas que son evidentes en tareas como la predicción en series temporales. Por ello, supone una buena práctica utilizarlo para detectar anomalías de la CEV en el presente trabajo investigativo.

4.3.8 Detección de fallas en aerogeneradores con arquitectura de transformadores

Chatterjee y Dethlefs (2020) realizaron un transformador DUAL para generar estrategias de mantenimiento correctivo, utilizando datos del sistema SCADA. El primer transformador identifica fallas basadas en las variables de entrada y su comportamiento para detectar alarmas; mientras que el segundo transformador realiza la selección según los códigos de alarma y crea estrategias en frases naturales contenidas en una plantilla. Los experimentos efectuados por Chatterjee y Dethlefs, a este modelo de transformador arrojan resultados del 96,75% en predecir alarmas y un 75,35% para elegir las estrategias de mantenimiento.

Las ventajas de la detección temprana de fallas en los aerogeneradores atraen a muchos investigadores, para beneficiarse del aumento de la confiabilidad y simplificar las tareas O&M de aerogeneradores. En este sentido, empleando los datos obtenidos por el sistema SCADA para cada aerogenerador, presentando una solución técnica y económicamente viable. Los conjuntos de datos recopilados por los sistemas SCADA son útiles para aplicar métodos de aprendizaje profundo basado en transformadores para predecir fallos en los aerogeneradores.

Con este argumento, este proyecto de investigación pretende desarrollar un modelo predictivo de detección de anomalías y varios componentes de los aerogeneradores de la CEV, usando datos de sistemas SCADA. Este modelo ayudará a desarrollar una estrategia mantenimiento predictivo efectivo para reducir las consecuencias de fallas prematuras y así reducir costos asociados a la O&M, y en consecuencia mejorar la disponibilidad y confiabilidad de la CEV.

4.3.9 Métodos de optimización modernos

En la actualidad, existen métodos que son capaces de realizar entrenamientos mucho más rápidos y alcanzando mejores resultados. A continuación, se presentan brevemente los más utilizados, puede encontrarse más detalles en Géron (2019):

- *Momentum*: Introduce el concepto de “momentum” en Gradiente Descendente, que genera un aumento en la distancia de cada paso si la dirección del gradiente es relativamente constante.
- *AdaGrad*: Introduce la tasa de aprendizaje adaptable, que acumula las dimensiones del gradiente, y disminuye la tasa de aprendizaje en las dimensiones más empinadas.
- *RMSProp*: Mejora problemas de AdaGrad al darle más importancia a los gradientes más recientes.

- *Adam*: Combina RMSProp con Momentum, generando una rápida convergencia con tasa de aprendizaje adaptable. Se ha convertido en el método de optimización estándar (Géron, 2019).

4.3.10 Métricas de evaluación

Estos métodos permiten evaluar el desempeño del modelo desarrollado, basándose en sus resultados, para determinar la eficacia del mismo. Existen varias métricas que se pueden utilizar para comparar el rendimiento del modelo:

4.3.10.1 Matriz de confusión en predicción

Es una tabla que resume el rendimiento de un modelo, donde un eje de la matriz de confusión es la etiqueta que predijo el modelo, mientras que la otra es la etiqueta real (Véase **Tabla 2**).

- Verdadero Positivo (TP *abreviatura del inglés*): La predicción de salida es correctamente clasificada como positiva.
- Falso Positivo (FP *abreviatura del inglés*): La predicción de salida es incorrectamente clasificada como positiva.
- Falso Negativo (FN *abreviatura del inglés*): La predicción de salida es correctamente clasificada como negativa.

Tabla 2. Matriz de Confusión en predicción.

Valores de predicción	Verdaderos Positivos TP	Falsos Positivos FP
	Falsos Negativos FN	N/A
	Valores reales	

Nota: Para predicción no se usa Falsos Positivos.

Se usa esta matriz para comparar entre modelos, observando que tan acertada es su predicción. El modelo base será el que predijo bien los TP.

4.3.10.2 Precisión (Precision)

Es la predicción clasificada correctamente como positiva entre el número de predicción total del modelo, de acuerdo a la **ecuación (4)** referida en Tharwat (2018).

$$precision = \frac{TP}{TP + FP} \quad (4)$$

4.3.10.3 Sensibilidad (Recall) o Tasa de verdadero positivo

Esta métrica se calcula como el número de verdaderos positivos dividido por el número total de verdaderos positivos y falsos negativos. Tharwat (2018) define que recall es la tasa de verdaderos positivos correctamente clasificadas del total de muestras positivas de acuerdo a la **ecuación (5)**.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

4.3.10.4 F1 Score

Corresponde a la media armónica de las medidas de precisión y recall, las cuales representan la exactitud de las predicciones positivas y la fracción de instancias positivas detectadas (Fuenza Méndez, 2020), como se muestra en la **ecuación (6)**.

$$F1 = 2 * \frac{precision * recall}{precision + recall} \quad (6)$$

Es la métrica preferida de clasificación, ya que elimina problemas de la exactitud.

4.3.11 Problemas típicos del aprendizaje automático

Existen algunos desafíos frecuentes que se afrontan, los cuales básicamente son:

4.3.11.1 Calidad de los datos

Los desafíos presentados son los valores faltantes, duplicados o incorrectos.

4.3.11.2 Dimensionalidad

Para Pertuz (2022), los modelos trabajan muy bien en un conjunto específico de atributos, pero si se crea un modelo con una gran cantidad de variables, puede que algunas de ellas no aporten verdadero valor, perdiendo rendimiento y exactitud.

4.3.11.3 Subajuste (Underfitting)

Este error ocurre cuando al modelo no se lo ha entrenado con los suficientes datos, por lo que no es capaz de relacionar las variables descriptivas y de destino (Pertuz, 2022). Los autores Zhang et al. (2019), argumentan que el ajuste insuficiente ocurre cuando un algoritmo carece de capacidad suficiente para aprender de la verdadera relación del conjunto de datos.

4.3.11.4 Sobreajuste (Overfitting)

Este asociado a modelos que se ajustan a los datos de entrenamiento, pero no generalizan bien, por ello no son capaces de predecir correctamente instancias nuevas (Pertuz, 2022). Es decir, un modelo con sobreajuste puede incluir información no deseada como ruido y valores atípicos. Bashir et al. (2020), definen que el sobreajuste ocurre cuando un algoritmo reduce el error mediante el aprendizaje de ejemplos con atributos ruidosos o irrelevantes, en lugar de su verdadera agrupación.

5. Metodología

En este apartado se presentan los materiales y métodos utilizados para desarrollar el modelo de predicción de fallas del módulo IGBT en los aerogeneradores de la CEV. Primero, se llevó a cabo el procesamiento de los datos e ingeniería de características de la base de datos del sistema SCADA. A continuación, se describió cómo se aplicó la técnica de aprendizaje profundo basada en transformadores. Finalmente, se detallaron los métodos utilizados para evaluar el rendimiento del modelo.

5.1 Área de estudio

El objeto de estudio, la CEV, se observa en la **Figura 1**, y se encuentra ubicada en un terreno complejo, específicamente en el cerro "Villonaco", a aproximadamente 14 km del centro de la ciudad de Loja, en la provincia de Loja. La CEV cuenta con 11 aerogeneradores de 1,5 MW de potencia nominal, modelo GW 70/1 500 de la marca GOLDWIND, con tecnología *Direct Drive* y generador de imanes permanentes. Las coordenadas de la CEV son UTM 693030 E y 9556476 N.

5.2 Características de los aerogeneradores en estudio.

Los datos técnicos de los aerogeneradores se observan en la **Tabla 3** a continuación.

Tabla 3. Datos técnicos del aerogenerador GOLDWIND GW 70/1 500

Datos técnicos	
Potencia	
Potencia nominal	1,5 MW
Velocidad de inicio	3 m/s
Velocidad de referencia del viento	11,8 m/s
Velocidad de corte	25 m/s
Rotor	
Diámetro	70 m
Área de barrido	3 850 m ²
Rango de velocidad	10,2 - 19 rpm
Palas	3
Tipo	LM 34P
Control de potencia	Control de Pitch
Generador	
Tipo	Generador síncrono de tecnología Direct Drive de imanes permanentes.
Voltaje nominal	690V
Velocidad de rotación nominal	19 rpm
Convertidor	
Tipo	Sistema modular IGBT
Rango del factor de potencia de salida	De -0,95 a +0,95
Corriente nominal de salida	1 397/1 255 A

Datos técnicos	
Sistema de orientación	
Tipo	3 motores de Inducción
Sistema de seguridad	
Frenado aerodinámico	Control de pitch independiente.
Freno mecánico	Sistema hidráulico (Bloqueo de rotor y freno de disco)
Torre	
Tipo	Torre de acero tubular
Altura de buje	65 m
Diseño estándar	IEC 1024-I
Sistema de control	
Tipo	SCADA

Fuente: Xinjiang Goldwind Science & Technology Co. (2011)

En el **Anexo 5**. Guía para ejecutar de manera efectiva los resultados de este trabajo Después de solicitar el código fuente a los responsables descritos en la sección de recomendaciones, recibirás una carpeta titulada 'Python Torres Cabrera Joel Stiven Torres'. Deberás cargar esta carpeta a tu Google Drive e instalar Google Colab. Es importante destacar que algunos procesos pueden requerir un mínimo de 30 GB de memoria RAM, para lo cual puedes utilizar Google Colab Pro.

Además, ten en cuenta que esta guía puede estar sujeta a cambios. En la carpeta compartida encontrarás un archivo de texto con la guía actualizada, así como posibles nuevas versiones.

Para el procesamiento, etiquetado y filtrado de los datos se deben seguir los siguientes pasos:

1. Para cargar los datos, dirígete a la carpeta 'Python Torres Cabrera Joel Stiven Torres/Dataset original/archivos m'. En esta ubicación deberás cargar los archivos obtenidos.

2014	U1	m140101.txt
2015	U2	m140102.txt
2016	U3	m140103.txt
2017	U4	m140104.txt
2018	U5	m140105.txt
2019	U6	m140106.txt
2020	U7	m140107.txt
2021	U8	m140108.txt
	U9	m140109.txt
	U10	m140110.txt
	U11	m140111.txt
		m140112.txt

Figura 208. Archivos de dataset de operación de los aerogeneradores.

2. Además, carga los archivos de alarmas del SCADA y el archivo de O&M en la ruta 'Python Torres Cabrera Joel Stiven Torres/Dataset original/fallas'. Aquí se muestran los archivos correspondientes.

X	Fallas 2014.xlsx
X	fallas 2015-2021.xlsx
X	Fallas 2015.xlsx
X	OT 2017-2021 (1).xlsm

Figura 209. Archivos de alarmas del SCADA y de mantenimiento de los aerogeneradores.

3. Después abrir el archivo "PROCESAMIENTO_DE_DATOS_COLAB_25GB RAM.ipynb" en Google Colab, asegúrate de cambiar la ruta de acceso para que apunte a la carpeta "Python Torres Cabrera Joel Stiven Torres". Esto garantizará que puedas acceder a los archivos necesarios para tu trabajo.

```
[ ] 1 #Aquí debe cambiar su ruta de archivo, por la ubicación de la carpeta "Python Torres Cabrera Joel Stiven Torres" descargada.  
2 Drive="/content/drive/MyDrive/Python Torres Cabrera Joel Stiven Torres"
```

Figura 210. Celda de código de ruta de archivos para procesamiento de datos.

4. Después de ejecutar todas las celdas, obtendrás el archivo "dataset_completo_V2_etiquetado.csv". Este archivo estará disponible en la carpeta "Python Torres Cabrera Joel Stiven Torres/Dataset original". Contendrá todos los datos extraídos y estará listo para su análisis posterior.

- Para la selección de características y visualización de variables y división de los datos.

5. Para proceder, abre el archivo "FSINGIENERIA_CHARACTERISTICAS_COLAB_25GB RAM.ipynb" y asegúrate de cambiar la ruta de acceso de la carpeta de la misma manera que se hizo en el paso 3 anterior. Esto garantizará que puedas trabajar con los datos de manera efectiva en Google Colab. Una vez ejecutado el cuaderno se descargarán datasets de entrenamiento.

- Entrenamiento y prueba del modelo Anomaly Transformer

6. Para entrenar el Modelo Anomaly Transformer modificado en este estudio, primero, acceda a la carpeta del modelo llamada "Anomaly-Transformer-modified_Torres-20240418T150717Z-001". Luego, abra el archivo "Ejecución_anomali_transformer.ipynb" y ejecute las celdas según sea necesario. Asegúrate de estar utilizando un entorno con GPU en Google Colab durante la ejecución de tu código.

7. Dentro de la carpeta, deberá modificar manualmente los parámetros descritos en el proyecto de investigación. Esto implica cambiar los parámetros (FeatSelect y FS) para el tipo de selección de características utilizado. Estos parámetros se modificarán en el archivo "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/utils/params.py".

8. Además, en el archivo "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-

modified_Torres/Anomaly-Transformer-main/solver.py", debe ajustar los parámetros (N, thresh, a, N2, thresh2, a2) según lo indicado en las tablas de resultados, dependiendo del aerogenerador y el tipo de selección de características.

9. En el archivo "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/scripts/WT.sh", se debe cambiar el umbral (r) de anomalías según se encuentre especificado en las tablas de resultados.

10. Luego, carga los archivos en formato numpy, los cuales resultan de la división de datos, a la carpeta "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/dataset/WT". Estos datos serán fundamentales para el entrenamiento y evaluación del modelo.

10. Finalmente, los resultados del sistema de predicción de fallas se cargarán en la carpeta "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/plots/WT".

- Entrenamiento y prueba del modelo TranAD

11. Para entrenar el Modelo TranAD modificado en este estudio, dirígete a la carpeta del modelo llamada "Anomaly-Transformer-modified_Torres-20240418T150717Z-001" y abre el archivo "TranAD_modified_Torres_COLAB 25GB RAM.ipynb". Asegúrate de tener al menos 30 GB de RAM disponible para el entrenamiento; puedes utilizar Google Colab Pro para esto.

12. Además, necesitarás ajustar los parámetros (r, N, thresh, a, N2, N3, threshEt2 y a2) según las especificaciones del estudio. Estos parámetros se deben cambiar en el archivo ipynb mencionado.

13. Luego, carga los archivos en formato numpy, los cuales resultan de la división de datos, a la carpeta "/Python Torres Cabrera Joel Stiven Torres/TranAD_Modified_Torres/TranAD/processed/WT_DATA". Estos datos serán fundamentales para el entrenamiento del modelo.

14. Finalmente, los resultados del sistema de predicción de fallas se cargarán en la carpeta “/Python Torres Cabrera Joel Stiven Torres/TranAD_Modified_Torres/TranAD/plots/TranAD_WT_DATA”.

Errores que podrías tener al ejecutar el código.

15. En caso de que surja algún error en la ejecución por desactualización, es fundamental asegurarse de actualizar las modificaciones que puedan haber ocurrido en las librerías utilizadas o en la versión de Python en Google Colab. Estos cambios pueden provocar incompatibilidades con algunas librerías. Mantener todo actualizado garantiza un funcionamiento óptimo del código y previene posibles errores relacionados con versiones desactualizadas. También puede contactarse con el equipo de responsables de este proyecto.

Anexo 6, se encuentran la especificación del fabricante respecto a la curva de potencia garantizada del aerogenerador.

5.3 Materiales y equipos

Los materiales y equipos utilizados se detallan a continuación:

5.3.1 Material bibliográfico

- Libros especializados en Aprendizaje Automático y Aprendizaje Profundo.
- Artículos científicos sobre temas relacionados con la detección de anomalías y la predicción de fallos en los aerogeneradores.
- Proyectos de grado relacionados con el campo del mantenimiento predictivo en los aerogeneradores.
- Artículos y libros centrados al procesamiento de datos e implementación de modelos para la arquitectura de transformadores.
- Datos del sistema SCADA de los aerogeneradores de la CEV.

5.3.2 Equipos Tecnológicos

- Computador

5.3.3 Recursos humanos

- Director de trabajo de titulación
- Asesor de trabajo de titulación
- Autor de trabajo de titulación

5.3.4 Lenguaje de programación

- Python

5.3.4.1 Librerías

- Pandas
- Numpy
- Matplotlib/Seaborn
- Torch
- Tensorflow
- Keras
- Scikit Learn
- Otros

5.3.5 Acuerdo de confidencialidad

Los datos del sistema SCADA de los aerogeneradores proporcionados son confidenciales y pertenecen a la Unidad de Negocios GENSUR EP. No es de acceso público. En el **Anexo 7**, se presenta el acuerdo de confidencialidad para el manejo de estos datos.

5.4 Procedimiento

5.4.1 Enfoque metodológico

El enfoque de la investigación es cuantitativo, debido a que se obtiene información de los diferentes sensores que son recolectados por el sistema SCADA de los aerogeneradores. Esto permite analizar y medir de forma objetiva los datos recolectados.

5.4.2 Tipo de diseño

Este diseño es de tipo longitudinal, ya que la base de datos SCADA contiene información de sensores a lo largo del tiempo, con una frecuencia de 10 minutos desde el 01 de enero del 2014 al 31 de octubre del 2021.

5.5 Metodología

En base a la revisión de bibliografía relacionada, se estableció una metodología para la implementación del proyecto. Esta metodología consta de las fases descritas en la **Figura 7**.

Estas fases incluyen el procesamiento y análisis de los datos, la implementación y prueba del modelo, la evaluación y mejora del modelo. Cada fase es esencial para el éxito del proyecto y ayuda a garantizar que se alcancen los objetivos específicos.



Figura 7. Metodología utilizada en la elaboración del proyecto.

5.6 Procedimiento y análisis de datos del objetivo 1

5.6.1 Procesamiento de datos

Para este apartado consulte la **Figura 8**, donde se presenta el procedimiento para el procesamiento de los datos.

5.6.1.1 Análisis de datos del sistema SCADA y procedimiento

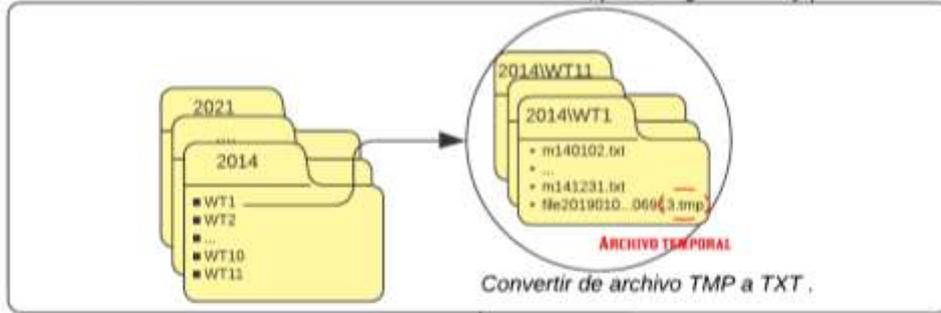
La base de datos recopilada del sistema SCADA se encontró en distintas carpetas con archivos en formato de texto (.txt) y temporales (.tmp). Se concatenaron en un solo archivo de formato .csv. Además, se solucionaron desafíos como archivos sin encabezados, archivos vacíos y valores faltantes.

5.6.1.2 Procesamiento de valores faltantes y perdidos.

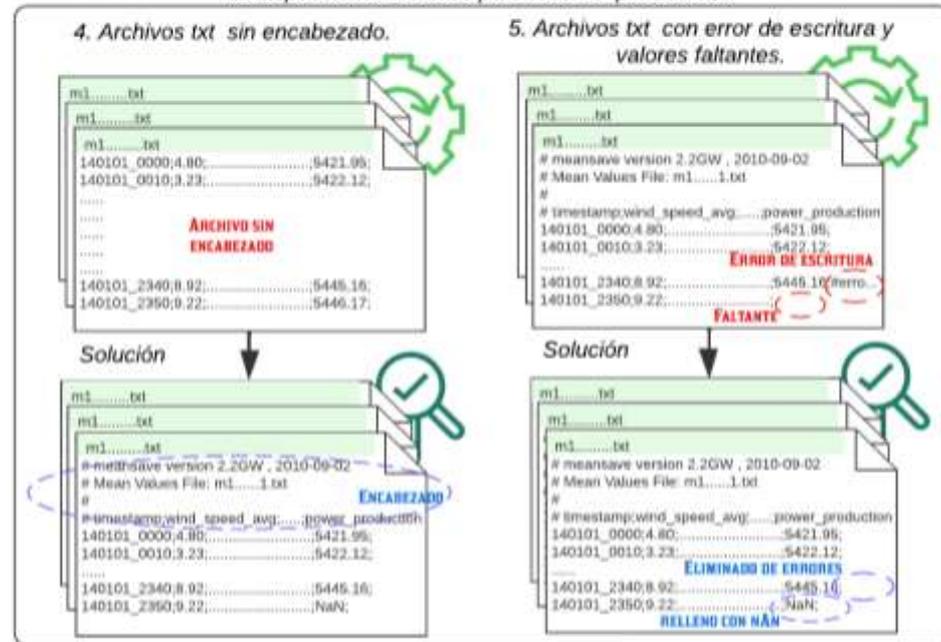
Los datos encontrados donde toda la fila no contenía ningún valor o faltaba algún dato, fueron eliminados. También se eliminaron datos en los que la columna "timestamp" contenía

números en lugar de fechas. Una vez eliminadas estas filas, se ordenaron los datos por fechas de forma consecutiva y por aerogenerador.

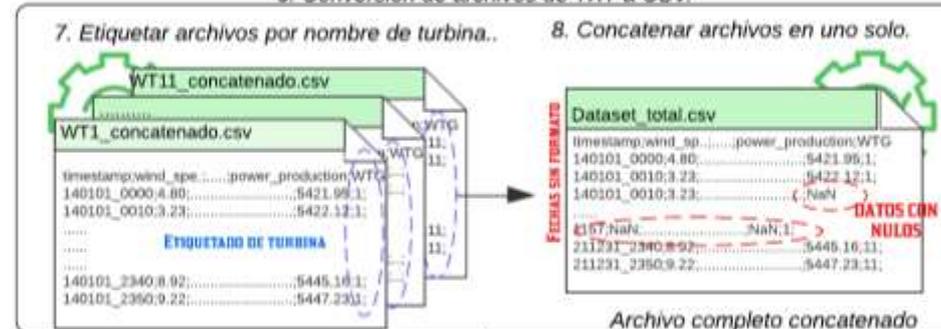
1. Los archivos se encuentran distribuidos en años, por aerogenerador y por día.



2. Inspeccionar archivos para resolver problemas.



6. Conversión de archivos de TXT a CSV.



8. Eliminación de filas con datos nulos (NaN) y formato a las fechas.

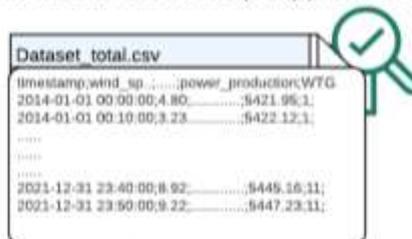


Figura 8. Procesado de datos.

5.6.2 Filtrado de información del conjunto de datos

De acuerdo con la información bibliográfica revisada sobre el procesado de datos en aerogeneradores, se debe tener en cuenta los siguientes aspectos:

- Potencia debe ser mayor a cero.
- Velocidad del viento mayor a velocidad de conexión.
- Velocidad del viento menor a la velocidad de corte.
- Temperatura mínima de todos los sensores debe ser mayor a 0°C
- Modo de operación en funcionamiento normal (sin restricciones de red).

5.6.3 Etiquetado de datos

Para etiquetar las fallas, se buscó en el registro de O&M las fallas que han ocurrido, en este caso centrándose en fallas del módulo IGBT. Este registro contiene información sobre las O&Ms realizados en todos los aerogeneradores desde el año 2017 hasta 2021. La forma en que está organizado el archivo de Excel se presenta en la **Tabla 4**.

Tabla 4. Formato de archivo de registro de operación y mantenimiento de la CEV.

MES	WTG	DESC.	TIPO	ÁREA	# OT	FECHA & HORA INICIO	FECHA & HORA FINAL	IND.	REP.	REP. REEMP.	CANT.
ene-2017	6	Instalación de Sensor Ultr.	CORR.	PROT.	714	13/1/2017 14:18	13/1/2017 16:40	SI	SI	SENSOR ULTR.	1
ene-2017	1	Falla de configuración PLC.	CORR.	PROT.	723	29/1/2017 3:05	29/1/2017 12:20	SI	SI	FUENTE DE PODER UPS	1
.....											
nov-17	5	Mantenimiento menor U5	PREV.	MEC.	731	13/2/2017 9:50	13/2/2017 16:46	SI	SI	PANTALLA TÁCTIL	1
dic-2017	4	FALLA EN CONVERT.	CORR.	ELÉ.	896	30/12/2017 8:45	30/12/2017 10:15	SI	SI	MÓDULO IGBT	1

Para el etiquetado de alarmas se procedió a recopilar información del archivo en formato Excel, donde se encuentran las alarmas producidas por el sistema SCADA desde el año 2014 al 2021. Esta información se presenta en la **Tabla 5**.

Tabla 5. Archivo Excel de alarmas dado por el sistema SCADA.

Fault start time	WTG	Fault End Time	Fault length (s)	Fault duration description(h)	Fault description	Maint.	wtypename
2018-02-18 11:43:10	S02	2018-02-18 11:46:57	228.00	0.063	(435)Error_converter not ready		GW1500
2018-02-28 17:32:45	S02	2018-02-28 17:35:14	149.00	0.041	(435)Error_converter not ready		GW1500
.....							
2021-10-04 15:09:39	S11	2018-03-04 16:32:07	4948.00	1.374	(217)Error_pitch blade angle comparing		GW1500
2021-10-04 16:35:51	S11	2018-03-04 16:36:09	19.00	0.005	(84) Error_Gen side capacitor fuse		GW1500

Para etiquetar donde ocurrió la falla o la alarma, se procedió a tomar la columna “Fecha y hora de inicio / Fault start time”, y se redondeó hacia abajo en múltiplos de 10 minutos para que coincidiera con la columna “timestamp” del archivo de datos operativos del sistema SCADA. En el archivo de O&M, se filtraron las fallas relacionadas con el módulo IGBT, y se aplicó un proceso similar con el archivo de alarmas del SCADA, pero este caso se filtraron todos los códigos de alarma relacionados con el componente en estudio.

Luego, se localizó la fecha y hora en que ocurrió la falla o alarma para etiquetarla. Si no se encontraba, se buscó una fecha hacia atrás iterativamente, debido a que existen fechas faltantes en los datos del sistema SCADA y también porque es posible que el componente fuera reemplazado tiempo después de haber ocurrido la falla. En la **Figura 9**, se presenta el flujograma utilizado para etiquetar fallas y alarmas relacionados al módulo IGBT.

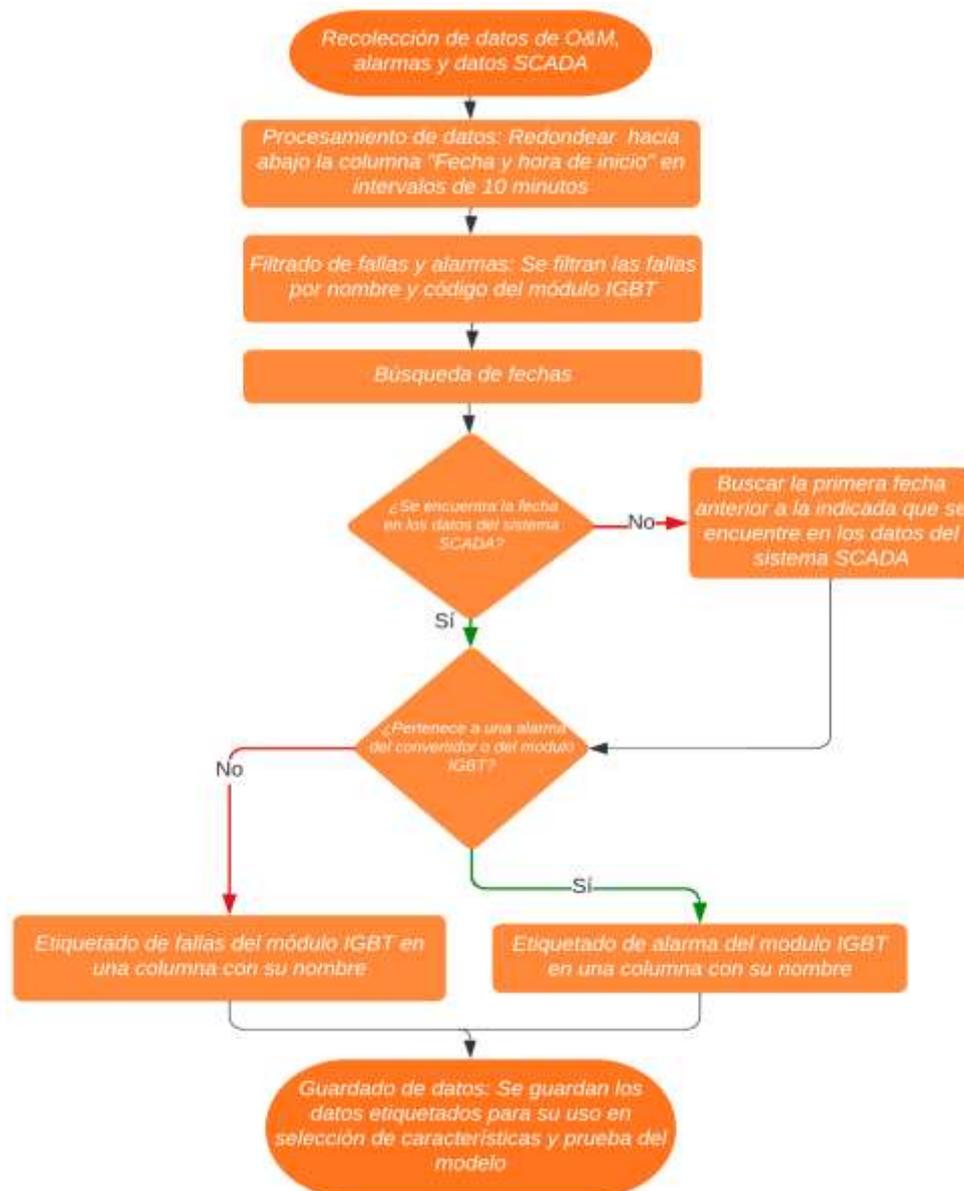


Figura 9. Flujograma de etiquetado de fallas y alarmas del módulo IGBT.

5.6.4 Selección de características

Se seleccionaron las variables mediante un enfoque combinado de métodos manuales y estadísticos. En primer lugar, se consideró la bibliografía relacionada para seleccionar métodos como el de correlación, eliminación recursiva de características, prueba de hipótesis estadística, selección basada en árboles y modelos lineales. Se realizó una selección de variables de entrada al modelo mediante prueba y error, ya que no existe una metodología establecida en la bibliografía para definir que método es más preciso. Estos métodos se encuentran detallados en la sección 4.2.3.1.

5.6.4.1 Escalado y normalización

Rashid y Batunlu (2021), Thoresen (2019) y Pertuz, (2022) destacan que los datos a diferentes escalas son difíciles de tratar y pueden provocar sesgos en el modelo de aprendizaje automático. Por esta razón, se realizó el escalado y normalización de las variables en un rango de 0 y 1, con el objetivo de evitar sesgos y facilitar el procesamiento de los datos.

5.6.4.2 División de los datos de entrenamiento, validación y prueba

En este caso los datos de entrenamiento, validación y prueba no se dividieron en base a lo común, ya que en base a literatura revisada para entrenar un modelo sin supervisión se debe entrenar en lo más posible con datos normales, por ello se dividió de forma única para cada aerogenerador donde se tomó para entrenamiento, datos desde 2014 hasta un año antes de que ocurra la falla, mientras que para prueba se encuentran los datos con las fallas que haya existido.

5.6.4.3 Análisis y comportamiento de variables SCADA

A continuación, se presenta algunas formas de representar el análisis y comportamiento de los datos SCADA:

- Gráficos de línea: Se utilizó gráficos de línea para mostrar cómo varían las variables medidas a lo largo del tiempo. Esto puede ser útil para observar patrones o tendencias en los datos.
- Gráficos de dispersión: Se utilizó gráficos de dispersión para visualizar las relaciones entre dos variables medidas. Esto puede ayudar a identificar patrones o tendencias en los datos.
- Box plots: Se utilizó gráficos de caja para mostrar la distribución de los datos y detectar valores atípicos.
- Mapas de calor: Utilizar mapas de calor para representar la distribución de los datos en una matriz.

- Gráficos de control: Utilizar gráficos de control para detectar patrones o tendencias en los datos.

5.6.5 Desarrollo y aplicación del modelo

Para el desarrollo del modelo de aprendizaje profundo basado en transformadores, se utilizó una metodología sencilla y eficaz. El objetivo de este modelo es predecir fallas con anticipación; es decir, antes que la falla cause la parada del aerogenerador. El modelo basado en transformadores para detección de anomalías es una arquitectura capaz de detectar patrones anómalos de las variables seleccionadas en el procesamiento de datos, permitiendo detectar posibles fallas antes de que ocurran. En la **Figura 10**, se puede comprender la secuencia que sigue el modelo de predicción para detectar patrones anómalos.



Figura 10. Flujograma de modelo de predicción de fallas.

En el mundo real, los datos de anomalías son escasos y difíciles de etiquetar al momento de detectar patrones anómalos. Por esta razón, el aprendizaje no supervisado es de importancia en este estudio. Los criterios de discriminación entre lo normal y lo anormal son útiles en esta tarea, ya que no se requieren datos de etiquetados en el entrenamiento. Existen algunos otros métodos como estimación de densidad y de agrupamiento, pero son difíciles de generalizar en aplicaciones reales.

La razón por la que la arquitectura del transformador es utilizada ampliamente es que puede expresar relaciones globales y de largo plazo de manera unificada. Cuando se aplica a

una serie temporal, los mapas de autoatención, son utilizados para detectar patrones y relaciones temporales en una serie de datos. Sin embargo, puede ser difícil encontrar patrones anómalos debido a la rareza de estos y al predominio de datos normales. Por lo tanto, se enfoca en asociaciones entre puntos de tiempo cercanos que probablemente contengan patrones similares de anomalías para asegurar continuidad. Este enfoque se conoce como asociación previa.

5.6.5.1 Implementación del primer modelo del Transformador Anomaly Transformer

Para implementar el primer modelo de transformador se usará la metodología usada por J. Xu et al. (2021). Este método ha sido escogido por su mecanismo de atención de anomalías basado en la observación clave de discordancia de asociaciones.

Debido a la limitación que presenta la arquitectura de transformador para la detección de anomalías, se ha renovado su estructura mediante la implementación de un mecanismo de atención de anomalías. El transformador de anomalías se compone de bloques de Anomaly-Attention y capas de feedforward, que se apilan alternativamente como se muestra en la **Figura 11**, lo que permite aprender posibles asociaciones a partir de las características profundas de capas múltiples en los datos. Esta estructura de apilamiento ayuda a mejorar la capacidad de detección de patrones anómalos en los datos. Expresado en fórmula en las ecuaciones (7) y (8):

$$Z^l = \text{Layer - Norm}(\text{Anomaly - Attention}(X^{l-1}) + X^{l-1}) \quad (7)$$

$$X^l = \text{Layer - Norm}(\text{Feed - Forward}(Z^l) + Z^l) \quad (8)$$

Donde $X^l \in \mathbb{R}^{N \times d_{model}}$, $l \in \{1, \dots, L\}$ denota la salida de la l -ésima capa con canales d_{model} . La entrada inicial $X^0 = \text{Embedding}(X)$ representa la serie sin procesar incrustada. $Z^l \in \mathbb{R}^{N \times d_{model}}$ es la representación oculta de la l -ésima capa. $\text{Anomaly - Attention}(\cdot)$ es para calcular la discrepancia de asociación.

En la **Figura 11** se muestra dos estructuras distintas que modelan la asociación previa y en serie. En la asociación previa, se utiliza un kernel gaussiano aprendible para reflejar el sesgo producido por las concentraciones cercanas en cada punto de tiempo. Por otro lado, la asociación en serie se refiere a los pesos de autoatención que se aprenden de la serie temporal original. Finalmente, se utiliza una estrategia de minimizar y maximizar entre estas ramas para aumentar la discriminación entre la discordancia de la asociación normal y anormal.

Este modelo se optimiza mediante la estrategia minimax para abordar el problema de la pérdida de reconstrucción. Se utiliza un mecanismo de gradiente de parada diseñado

específicamente (representado por flechas grises) para restringir las asociaciones previas y en serie, lo que permite una mayor diferenciación entre las discrepancias de asociación y, por lo tanto, una mejor capacidad de discriminación.

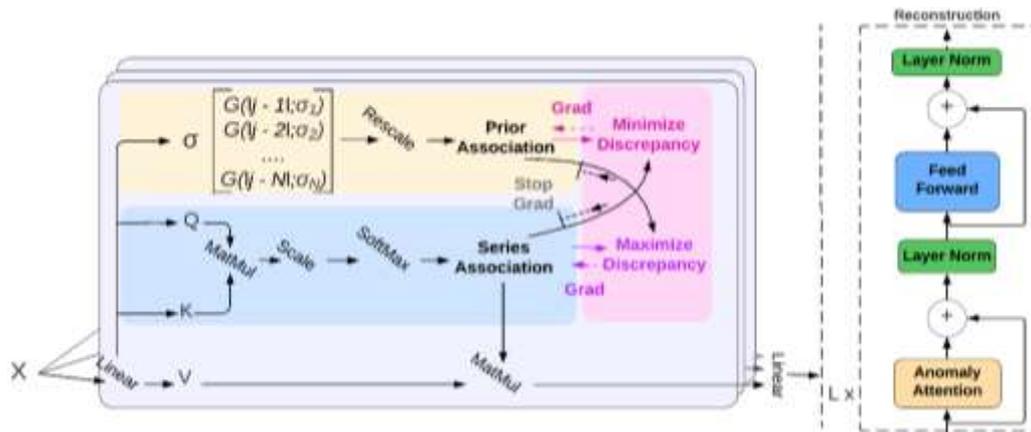


Figura 11. Metodología de Anomaly Transformer.
Fuente: Adaptado de J. Xu et al. (2021)

5.6.5.1.1 Atención de anomalías

El modelo de detección de anomalías utiliza dos estructuras para capturar las asociaciones temporales y previas en los datos. La primera estructura analiza la distancia temporal entre los puntos de datos mediante un kernel gaussiano aprendible, mientras que la segunda se centra en identificar patrones anómalos en segmentos de datos de diferentes longitudes. Estas dos estructuras trabajan juntas para mejorar la capacidad del modelo para detectar patrones anómalos en los datos. La asociación en serie aprende directamente de los datos sin procesar, adaptándose automáticamente a las asociaciones más efectivas. Esto preserva la relación temporal entre puntos y produce una representación más informativa de la serie temporal. La forma en que se expresa esta información se presenta en las fórmulas matemáticas de las ecuaciones (9), (10), (11) y (12).

$$Initialization: Q, K, V, \sigma = X^{l-1}W_Q^l, X^{l-1}W_K^l, X^{l-1}W_V^l, X^{l-1}W_\sigma^l \quad (9)$$

$$Prior_{Association}: P^l = Rescale \left(\left[\frac{1}{\sqrt{2\pi}\sigma_i} \exp \left(-\frac{|j-i|^2}{2\sigma_i^2} \right) \right]_{i,j \in \{1,..,N\}} \right) \quad (10)$$

$$Series_{Association}: S^l = Softmax \left(\frac{QK^T}{\sqrt{d_{model}}} \right) \quad (11)$$

$$Reconstruction: \hat{Z}^l = S^l V, \quad (12)$$

Donde $Q, K, V \in R^{N \times d_{model}}, \sigma \in R^{N \times 1}$ representa la consulta, la clave, el valor de la autoatención y la escala aprendida, respectivamente. $W_Q^l, W_K^l, W_V^l \in R^{d_{model} \times d_{model}}, W_\sigma^l \in R^{d_{model} \times 1}$ representan las matrices de parámetros para Q, K, V, σ en la l -ésima capa respectivamente. La asociación previa $P^l \in R^{N \times N}$ se genera en base a la escala aprendida $\sigma \in R^{N \times 1}$ y el i -ésimo elemento σ_i corresponde al i -ésimo punto de tiempo. Concretamente, para el i -ésimo punto de tiempo, su peso de asociación con el j -ésimo punto se calcula mediante el kernel gaussiano $G(|j - i|; \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{|j-i|^2}{2\sigma_i^2}\right)$ con la distancia $|j - i|$. Además, usó $Rescale(\cdot)$ para transformar los pesos de asociación en distribuciones discretas P^l dividiendo la suma de las filas. $S^l \in R^{N \times N}$ denota las asociaciones en serie. $Softmax(\cdot)$ normaliza el mapa de atención a lo largo de la última dimensión.

Así, cada fila de S^l forma una distribución discreta. $\hat{Z}^l \in R^{N \times d_{model}}$ es la representación oculta después de Anomaly-Attention en la l -ésima capa. Se usa $Anomaly-Attention(\cdot)$ para resumir las ecuaciones (9), (10), (11) y (12). En la versión de múltiples cabezas que usa la escala aprendida es $\sigma \in R^{N \times h}$ para h cabezas. $Q_m, K_m, V_m \in R^{N \times \frac{d_{model}}{h}}$ denotan la consulta, la clave y el valor de la cabeza m -ésima respectivamente. El bloque concatena las salidas $\left\{ \hat{Z}_m^l \in R^{N \times \frac{d_{model}}{h}} \right\}_{1 \leq m \leq h}$ de varios cabezales y obtiene el resultado final $\hat{Z}_m^l \in R^{N \times d_{model}}$.

A continuación, se presenta el pseudocódigo utilizado por J. Xu et al. (2021) para atención de anomalías en el **Algoritmo 1**.

Algorithm 1 Anomaly-Attention Mechanism (multi-head version).

Input: $X \in R^{N \times d_{model}}$: input; $D = ((j - i)^2)_{i,j \in \{1, \dots, N\}} \in R^{N \times N} N \times N$: relative distance matrix
Layer params: MLP_{input} : linear projector for input; MLP_{output} : linear projector for output

1: $Q, K, V, \sigma = \text{Split}(MLP_{input}(X), \text{dim}=1)$ 2: for $(Q_m, K_m, V_m, \sigma_m)$ in (Q, K, V, σ) : 3: $\sigma_m = \text{Broadcast}(\sigma_m, \text{dim}=1)$ 4: $P_m = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-\frac{D}{2\sigma_m^2}\right)$ 5: $P_m = P_m / \text{Broadcast}(\text{Sum}(P_m, \text{dim}=1))$ 6: $S_m = \text{Softmax}\left(\sqrt{\frac{h}{d_{model}}}\frac{Q_m K_m^T}{d_{model}}\right)^n$ 7: $\hat{Z}_m = S_m V_m$ 8: $\hat{Z} = MLP_{output}(\text{Concat}([Z_1, \dots, \hat{Z}_h], \text{dim}=1))$ 9: Return \hat{Z}	$Q, K, V \in R^{N \times d_{model}}, \sigma \in R^{N \times h}$ $Q_m, K_m, V_m \in R^{N \times \frac{d_{model}}{h}}, \sigma_m \in R^{N \times 1}$ $\sigma_m \in R^{N \times N}$ $P_m \in R^{N \times N}$ Rescaled $P_m \in R^{N \times N}$ $S_m \in R^{N \times N}$ $\hat{Z}_m \in R^{N \times \frac{d_{model}}{h}}$ $\hat{Z} \in R^{N \times d_{model}}$ Keep the P_m and $S_m, m = 1, \dots, h$
--	---

5.6.5.1.2 Discrepancia de asociación

La diferencia entre la asociación previa y la asociación en serie se cuantifica mediante una medida simétrica de divergencia KL. En la ecuación (13), se representa cuán diferentes son dos distribuciones promedio, que se obtiene a partir de la variación en la asociación a través de múltiples capas.

$$AssDis(P, S; X) = \left[\frac{1}{L} \sum_{l=1}^L (KL(P_i^l, : || S_i^l, :) + KL(S_i^l, : || P_i^l, :)) \right]_{i=1, \dots, N} \quad (13)$$

Donde $KL(\cdot, || \cdot)$ es la divergencia KL calculada entre dos distribuciones discretas correspondientes a cada fila de P^l y S^l . $AssDis(P, S; X) \in R^{N \times 1}$ es la discrepancia de asociación puntual de X con respecto a la asociación previa P y la asociación en serie S de múltiples capas. El i -ésimo elemento de los resultados corresponde al i -ésimo punto de tiempo de X . $AssDis$ es más pequeño en datos anormales que datos normales. A continuación, se presenta el pseudocódigo utilizado por J. Xu et al. (2021) para la discrepancia de asociación en el **Algoritmo 2**.

Algorithm 2 Association Discrepancy $AssDis(P, S; X)$ Calculation (multi-head version).

Input: time series length N ; layers number L ; heads number h ; prior-association $P_{all} \in$

$R^{L \times h \times N \times N}$; series-association $S_{all} \in R^{L \times h \times N \times N}$;

1: $P' = \text{Mean}(P, \text{dim}=1)$

$P' \in R^{L \times N \times N}$

2: $S' = \text{Mean}(S, \text{dim}=1)$

$S' \in R^{L \times N \times N}$

3: $R' = \text{KL}((P', S'), \text{dim}=-1) + \text{KL}((S', P'), \text{dim}=-1)$

$R' \in R^{L \times N}$

4: $R = \text{Mean}(R', \text{dim}=0)$.

$R \in R^{N \times 1}$

5: **Return** R Represent the association discrepancy of each time point

5.6.5.1.3 Aprendizaje de asociación min-max

Para detectar patrones anómalos, se emplea una técnica de medición de divergencia que compara la asociación previa con la asociación secuencial. Para detectar patrones anómalos, se emplea una técnica de medición de divergencia que compara la asociación previa con la asociación secuencial. El modelo se optimiza mediante una técnica que utiliza una pérdida de discrepancia para aumentar la detección de las anomalías, lo que permite identificarlas de manera más efectiva incluso en regiones no adyacentes de la serie temporal. La función de pérdida es una combinación de la capacidad del modelo para reconstruir los datos originales y su habilidad para detectar patrones anómalos. Esto se logra mediante la suma de la pérdida de reconstrucción y el desequilibrio de asociación en el modelo, así como la ecuación (14).

$$L_{Total}(\hat{X}, P, S, \lambda; X) = \left\| X - \hat{X} \right\|_F^2 - \lambda \times \left\| AssDis(P, S; X) \right\|_1 \quad (14)$$

Donde $\hat{X} \in R^{N \times d}$ denota la reconstrucción de X . $\|\cdot\|_F, \|\cdot\|_k$ indica la norma de Frobenius y la norma k . λ es para equilibrar los términos de pérdida. Cuando $\lambda > 0$, se optimiza es para aumentar la discrepancia de asociación. J. Xu et al. (2021) proponen una estrategia minimax para hacer que la discrepancia de asociación sea más distinguible.

5.6.5.1.4 Estrategia de mini-max

La función de pérdida del modelo de detección de anomalías se compone de dos partes: la pérdida de reconstrucción y el desajuste de asociación. El desajuste de asociación se refiere a la discrepancia entre la asociación previa, que se basa en un kernel gaussiano, y la asociación secuencial, que se basa en la serie temporal sin procesar. El objetivo es maximizar esta discrepancia para encontrar patrones anómalos, pero si se maximiza directamente, se reduciría el parámetro de escala del kernel gaussiano, lo cual no tiene sentido. Por lo tanto, se emplea una estrategia minimax en la que se busca un equilibrio entre minimizar la discrepancia entre la asociación previa y la asociación secuencial y maximizar la discrepancia entre lo normal y lo anormal. Esto permite adaptar la adaptación previa a diferentes patrones temporales. Se utiliza un enfoque minimax para equilibrar ambos aspectos y asegurar que la adaptación a diferentes patrones temporales sea efectiva.

Se emplea una estrategia de minimizar y maximizar simultáneamente la asociación previa y la asociación en serie para mejorar la detección de patrones anómalos. La fase de minimización se concentra en que las asociaciones previas se ajustan lo más posible a las asociaciones aprendidas de la serie sin procesar, lo que permite una adaptación a diferentes patrones temporales. En cambio, la fase de maximización se enfoca en incrementar la diferencia entre las asociaciones, lo que hace que la asociación en serie se centre más en patrones anómalos. Durante la fase de minimización, la asociación previa se ajusta para minimizar la discrepancia de asociación dentro de la familia de distribución derivada del kernel gaussiano. En la fase de maximización, la asociación en serie maximiza la discrepancia de asociación bajo la pérdida de reconstrucción. Así, integrando la pérdida de reconstrucción, las funciones de pérdida en las dos fases son:

$$\begin{aligned} \text{Minimize Phase: } L_{Total}(\hat{X}, P, S_{detach}, -\lambda; X) \\ \text{Maximize Phase: } L_{Total}(\hat{X}, P_{detach}, S, \lambda; X) \end{aligned} \quad (15)$$

Donde $\lambda > 0$ y $*_{detach}$ significa detener la retropropagación del gradiente de asociación (Véase **Figura 11**). Como P se aproxima a S_{detach} en la fase de minimización, la fase de maximización, se aplica una restricción más rigurosa a la asociación en serie, lo que hace que

los puntos de tiempo presten mayor atención al área no adyacente. Dado que, bajo la pérdida de reconstrucción, es mucho más difícil para las anomalías lograr esto en comparación con los puntos de tiempo normales, se amplifica la capacidad de distinguir entre anomalías y puntos de tiempo normales mediante la discrepancia de asociación.

5.6.5.1.5 Criterios de anomalía basados en relaciones

La detección de patrones anómalos en la serie temporal se logra mediante la combinación de la pérdida de reconstrucción y la medida de divergencia entre las asociaciones previas y secuenciales. Esta combinación permite identificar de manera efectiva patrones anómalos en la serie temporal. Para obtener el puntaje de anormalidad, se utiliza la **ecuación (16)** que involucra la combinación de la pérdida de reconstrucción y la discordancia de asociación normalizada.

$$Anomaly_{Score}(X) = Softmax(-AssDis(P, S; X)) \odot [||X_{i,:} - \hat{X}_i||_2^2]_{i=1,\dots,N} \quad (16)$$

Donde \odot es la multiplicación de elemento por elemento, $Anomaly_{Score}(X) \in \mathbb{R}^{N \times 1}$ denota el criterio de anomalía puntual de X . Además, el modelo de transformador de anomalías puede adaptarse a distribuciones de datos cambiantes a lo largo del tiempo, lo que lo hace adecuado para detectar anomalías en los datos SCADA de las turbinas eólicas. Finalmente, se presenta el pseudocódigo utilizado por J. Xu et al. (2021) para la asociación basada en criterio en el **Algoritmo 3**.

Algorithm 3 Association-based Criterion AnomalyScore(X) Calculation

Input: time series length N ; input time series $X \in \mathbb{R}^{N \times d}$; reconstruction time series $\hat{X} \in \mathbb{R}^{N \times d}$; association discrepancy $AssDis(P, S; X) \in \mathbb{R}^{N \times 1}$;

1: $C_{AD} = Softmax(-AssDis(P, S; X), dim=0)$

$C_{AD} \in \mathbb{R}^{N \times 1}$

2: $C_{Recon} = Mean((X - \hat{X})^2, dim=1)$

$C_{Recon} \in \mathbb{R}^{N \times 1}$

3: $C = C_{AD} \times C_{Recon}$

$C \in \mathbb{R}^{N \times 1}$

4: **Return** C

Anomaly score for each time point

5.6.5.2 Implementación del segundo modelo del Transformador TranAD

En el segundo modelo de transformador se va utilizar la metodología usada por Tuli et al. (2022), donde utiliza una refactorización inteligente de la arquitectura del transformador para la tarea de detección de anomalías en datos de series temporales. Al igual que otros modelos encoder-decoder, en un transformador, una secuencia de entrada se somete a varias transformaciones basadas en la atención. La arquitectura de la red neuronal en TranAD se muestra en la **Figura 12**. El codificador tiene la tarea de codificar la secuencia completa hasta el instante de tiempo actual C , utilizando un puntaje de enfoque que se explicará con más detalle más adelante. A continuación, el codificador de ventana utiliza esta información para generar

una representación codificada de la ventana de entrada W , que luego se pasa a dos decodificadores para crear su reconstrucción.

Para procesar una secuencia multivariada como W o C , primero se transforma en una forma matricial con modalidad m . Luego se utiliza una técnica de atención del producto escalado-puntos con tres matrices: Q (consulta), K (clave) y V (valor).

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{m}}\right)V \quad (17)$$

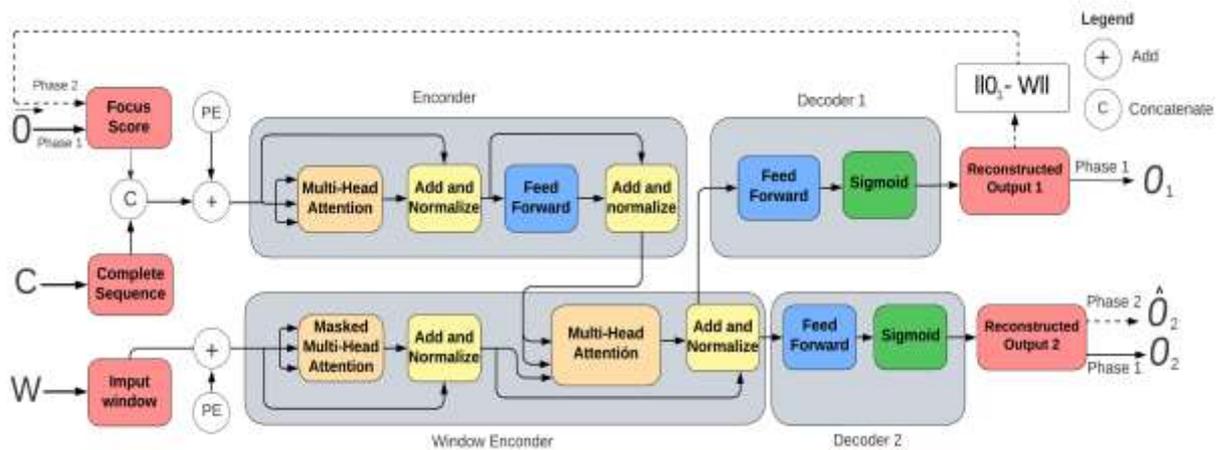


Figura 12. Metodología de TranAD.

Fuente: Adaptado de Tuli et al. (2022)

La función softmax es útil para asignar pesos a cada elemento en la matriz V y así reducir su tamaño para facilitar su procesamiento en las operaciones posteriores del modelo. Además, se utiliza un método de atención escalado que ajusta los pesos mediante un término de \sqrt{m} para estabilizar el entrenamiento del modelo. En cuanto a las matrices de entrada Q , K y V , se utiliza la técnica de atención múltiple con varios cabezales, lo que permite analizar la matriz de entrada desde diferentes perspectivas (determinadas por el número de cabezales h) para obtener una representación más completa del patrón temporal, así obtener Q_i , K_i y V_i para $i \in \{1, \dots, h\}$ y luego se aplica scaled-dot product attention como se observa en la ecuación (18):

$$MultiHeadAtt(Q, K, V) = Concat(H_1, \dots, H_h) \quad (18)$$

donde $H_i = Attention(Q_i, K_i, V_i)$

Un GAN (Generative Adversarial Network), que significa Red Generativa Adversarial, es un modelo de aprendizaje profundo que se utiliza para generar imágenes, audio, texto y otros tipos de datos. Está compuesto por dos redes neuronales en competencia: un generador y un discriminador. El generador produce datos artificiales mientras que el discriminador trata de identificar si los datos que recibe son reales o falsos. Por lo tanto, para la detección de

anomalías se utiliza un modelo eficiente en tiempo que se basa en un método de entrenamiento adversario de estilo GAN. Este modelo incluye dos codificadores transformadores y dos decodificadores. La inferencia del modelo se realiza en dos etapas. En primer lugar, se garantizaron los valores de W y C como entrada, junto con una puntuación de enfoque inicial F , que es una matriz de ceros con la misma dimensión que W . Luego, se amplía F para que tenga la misma dimensión que W , agregando rellenos de cero apropiados, y se concatena con W . Después de esto, se aplica la codificación de posición y se obtiene la entrada para el primer codificador, que se representa como I_1 . El primer codificador realiza las siguientes operaciones de la ecuación (19).

$$\begin{aligned} I_1^1 &= \text{LayerNorm}(I_1 + \text{MultiHeadAtt}(I_1, I_1, I_1)) \\ I_1^2 &= \text{LayerNorm}(I_1^1 + \text{FeedForward}(I_1^1)) \end{aligned} \quad (19)$$

Donde $\text{MultiHeadAtt}(I_1, I_1, I_1)$, denota la operación de múltiples cabezas para la matriz de entrada I_1 y “+” denota la suma de matrices. Las operaciones anteriores generan pesos de atención que capturan tendencias temporales dentro de las ventanas de series de tiempo y la secuencia completa de entrada. Además, el modelo incluye un codificador de ventana que se enfoca en la posición actual en el lugar de futuras marcas de tiempo para mejorar el tiempo de entrenamiento y evitar problemas en la inferencia. Estas operaciones permiten al modelo analizar múltiples lotes de ventanas de series de tiempo simultáneamente y mejorar la eficiencia del entrenamiento. El codificador de ventana realiza las siguientes operaciones de la ecuación (20).

$$\begin{aligned} I_2^1 &= \text{Mask}(\text{MultiHeadAtt}(I_2, I_2, I_2)) \\ I_2^2 &= \text{LayerNorm}(I_2 + I_2^1) \\ I_2^3 &= \text{LayerNorm}(I_2^2 + \text{MultiHeadAtt}(I_2^2, I_2^2, I_2^2)) \end{aligned} \quad (20)$$

El modelo utiliza la codificación de la secuencia completa, I_2 , como valores y claves para que el codificador de ventana realice la operación de atención con la ventana de entrada codificada como matriz de consulta. La idea detrás de las operaciones en la ecuación (19) es similar a la de la ecuación (20), pero en este caso se aplica un enmascaramiento de la entrada de la ventana para ocultar las marcas de tiempo futuro dentro del mismo lote de entrada. Al proporcionar al modelo la secuencia completa de entrada hasta el momento t , permite al modelo capturar y utilizar un contexto más amplio en comparación con uno que está limitado y acotado. Finalmente, se utiliza dos decodificadores idénticos que realizan la operación de la ecuación (21).

$$O_i = \text{Sigmoid}(\text{FeedForward}(I_2^3)) \quad (21)$$

Donde $i \in \{1,2\}$ corresponde al primer y segundo decodificador respectivamente. Así, este modelo de transformador toma la entrada C y W para generar las salidas O_1 y O_2 .

En **algoritmo 4**, se presenta el entrenamiento de TranAD desarrollado por Tuli et al. (2022).

Algorithm 4 The Tran AD training algorithm

Require:
Encoder E , Decoders D_1 and D_2
Dataset used for training W
Evolutionary hyperparameter ϵ
Iteration limit N
1: Initialize weights E, D_1, D_2
2: $n \leftarrow 0$
3: **do**
4: **for**($t = 1$ to T)
5: $O_1, O_2 \leftarrow D_1(E(W_t, \vec{0})), D_2(E(W_t, \vec{0}))$
6: $\hat{O}_2 \leftarrow D_2(E(W_t, \|O_1 - W_t\|_2))$
7: $L_1 = \epsilon^{-n} \|O_1 - W_t\|_2 + (1 - \epsilon^{-n}) \|\hat{O}_2 - W_t\|_2$
8: $L_2 = \epsilon^{-n} \|O_2 - W_t\|_2 - (1 - \epsilon^{-n}) \|\hat{O}_2 - W_t\|_2$
9: Update weights of E, D_1, D_2 using L_1, L_2
10: $n \leftarrow n + 1$
11: Meta-Learn weights E, D_1, D_2 using a random batch
12: **while** $n < N$

5.6.5.2.1 Entrenamiento adversarial de dos fases fuera de línea

En la fase 1, **entrada de reconstrucción**, el modelo Transformador se emplea para predecir la reconstrucción de cada ventana de la serie temporal de entrada, utilizando una arquitectura de codificador-decodificador para cada marca de tiempo. A diferencia de los modelos convencionales, el enfoque de inferencia es autorregresivo en dos fases para mejorar la detección de anomalías. La primera fase se centra en generar una aproximación de la entrada, mientras que la segunda fase se condiciona a las desviaciones en la primera fase. Esto permite al modelo capturar las tendencias temporales y detectar anomalías en áreas donde las desviaciones son más altas. Por lo que, en la primera etapa, los codificadores convierten la ventana de entrada $W \in R^{K \times m}$ (con puntuación de enfoque $F = [0]_{K \times m}$) a una representación latente comprimida I_2^3 utilizando la atención basada en el contexto como en un modelo transformador común. Esta representación comprimida luego se convierte para generar las salidas O_1 y O_2 .

En la fase 2, **enfoco de reconstrucción de entrada**, el modelo implementa una técnica de inferencia autorregresiva de dos fases para predecir la reconstrucción de cada ventana de

serie temporal de entrada. Durante la primera fase, se genera una puntuación de enfoque que indica las desviaciones entre la salida reconstruida y la entrada de datos. En la segunda fase, se utiliza esta puntuación para ajustar los pesos de atención y enfocar la red neuronal en subsecuencias de entrada específica, lo que ayuda a extraer tendencias temporales a corto plazo. Este enfoque presenta varios beneficios, como la mejora de la detección de anomalías, la prevención de falsos positivos y una mayor robustez en el rendimiento del modelo para diferentes secuencias de entrada.

En la fase 3, **el objetivo de entrenamiento en evolución**, en esta fase, el objetivo es lograr un entrenamiento estable. Por ello, se ha desarrollado un modelo de entrenamiento adversarial que utiliza las salidas de dos decodificadores diferentes. En la etapa inicial, ambos decodificadores se centran en la tarea de reconstruir la ventana de la serie temporal de entrada de forma individual. Se define la pérdida de reconstrucción para cada decodificador mediante el uso de la norma L_2 utilizando las salidas de la primera fase, como se describen en la ecuación (22).

$$L_1 = ||O_1 - W||_2 \quad (22)$$

$$L_2 = ||O_2 - W||_2$$

En la fase 2 del entrenamiento del modelo, se busca la estabilidad utilizando una técnica adversarial que involucra la salida de dos decodificadores diferentes. El objetivo del segundo decodificador es distinguir entre la ventana de entrada y la reconstrucción generada por el primer decodificador en la fase 1, maximizando la diferencia entre la salida del segundo decodificador y la entrada original $||O_2 - W||_2$. Por otro lado, el primer decodificador trata de engañar al segundo decodificador arrojando una puntuación de enfoque degenerado (un vector cero) cuando se reconstruye perfectamente la entrada (es decir, cuando la salida del primer decodificador es igual a la entrada original $O_1 = W$). Esto empuja al decodificador 2 a generar la misma salida que O_2 , lo que busca emparejar la entrada en la fase 1. Esto significa que el objetivo de entrenamiento es cómo la ecuación (23).

$$\min_{\text{Decoder1}} \max_{\text{Decoder2}} ||\hat{O}_2 - W||_2 \quad (23)$$

El objetivo del primer decodificador es reducir al mínimo el error de reconstrucción de la salida autocondicionada, mientras que el objetivo del segundo decodificador es maximizar ese mismo error. Esto se logra mediante el uso de una pérdida específica en la ecuación (24), donde se busca equilibrar estos objetivos opuestos.

$$L_1 = +||\hat{O}_2 - W||_2 \quad (24)$$

$$L_2 = -\|\hat{O}_2 - W\|_2$$

Una vez definidas las funciones de pérdida para cada fase, es necesario calcular la pérdida total acumulada para cada decodificador. Para ello, se emplea una función de pérdida evolutiva que combina las funciones de pérdida de reconstrucción y adversaria de ambas fases como se presenta en la ecuación (25).

$$\begin{aligned} L_1 &= \varepsilon^{-n}\|O_1 - W\|_2 + (1 - \varepsilon^{-n})\|\hat{O}_2 - W\|_2 \\ L_2 &= \varepsilon^{-n}\|O_2 - W\|_2 - (1 - \varepsilon^{-n})\|\hat{O}_2 - W\|_2 \end{aligned} \quad (25)$$

Donde n es el número de épocas del entrenamiento, y ε es un parámetro de entrenamiento cercano a uno (líneas 7-8 de Algoritmo 4). Al principio del proceso de entrenamiento, se reduce la importancia de la pérdida adversaria debido a que la segunda fase puede ser poco confiable para indicar reconstrucciones alejadas de la secuencia de entrada. Con el fin de evitar cualquier posible desestabilización en el entrenamiento del modelo, se le asigna un peso bajo a la pérdida adversaria. A medida que las reconstrucciones se acercan a las ventanas de entrada y los puntos de enfoque se vuelven más precisos, el peso de la pérdida adversaria se incrementa gradualmente. Para el proceso de entrenamiento, utilizamos pesos en forma exponencial con una pequeña constante positiva, como es común en las curvas de pérdida de redes neuronales. En consecuencia, se utilizan pesos de la forma ε^{-n} . Durante el proceso de entrenamiento, no se asume que los datos de la serie temporal están disponibles de manera secuencial, por lo que se pueden dividir los datos en pares (W, C) y utilizar lotes de entrada en el modelo. El uso de la técnica de atención multicabezal con máscaras permite la ejecución paralela a través de varios lotes, lo que acelera el proceso de entrenamiento. Además, esta técnica permite al modelo enfocarse en diferentes aspectos de los datos en cada lote de entrada, lo que puede mejorar su precisión y capacidad de generalización.

Se adquirió la técnica de Meta Learning, que consiste en un tipo de aprendizaje por refuerzo desarrollado por Finn et al. (2017), para mejorar el rendimiento del modelo TranAD en la detección de patrones y tendencias temporales en la serie temporal de entrenamiento, a pesar de la limitación de datos. Durante el entrenamiento, se realizan actualizaciones de gradiente para los pesos de la red neuronal (representados por θ), lo que ayuda al modelo a mejorar su capacidad de generalización y adaptación a los datos de entrada. Además, el aprendizaje por refuerzo permite al modelo aprender de sus errores y mejorar su desempeño a través de la iteración. Se expresa mediante la notación matemática de la ecuación (26).

$$\theta' \leftarrow \theta - \alpha \nabla_{\theta} L(f(\theta)) \quad (26)$$

Donde α , $f(\cdot)$ y $L(\cdot)$ son la tasa de aprendizaje, la representación abstracta de la red neuronal y la función de pérdida, respectivamente. Ahora, al final de cada época, realizamos un paso de meta-aprendizaje como se describe en la ecuación correspondiente. Este proceso de meta-aprendizaje ayuda al modelo a actualizar sus parámetros utilizando la información aprendida en la época anterior y mejorar su capacidad de generalización. Es importante elegir cuidadosamente una tasa de aprendizaje adecuada α y una función de pérdida $L(\cdot)$ apropiada para evitar el sobreajuste y asegurarse de que el modelo aprenda de manera efectiva. Ahora al final de cada época se usa la expresión matemática (27).

$$\theta \leftarrow \theta - \beta \nabla_{\theta} L(f(\theta')) \quad (27)$$

La meta-optimización se realiza con un tamaño de paso meta β , sobre los pesos del modelo θ , donde el objetivo se evalúa utilizando los pesos actualizados θ' . Trabajos previos como el que utilizó Finn et al. (2017) han demostrado que esto permite entrenar modelos rápidamente con datos limitados. Esta ecuación se muestra en la línea 11 del algoritmo 4.

5.6.5.2.1 *Detección de anomalías*

Ahora describimos el procedimiento de inferencia utilizando el modelo de transformación entrenado (resumido en el Algoritmo 5). Para datos no vistos $(\widehat{W}, \widehat{C})$ la puntuación de anomalía se define como:

$$s = \frac{1}{2} \|O_1 - \widehat{W}\|_2 + \frac{1}{2} \|\widehat{O}_2 - \widehat{W}\|_2 \quad (28)$$

La inferencia en el momento de la prueba se ejecuta nuevamente en dos fases y, por lo tanto, obtenemos un solo par de reconstrucción (O_1, \widehat{O}_2) (líneas 2 y 3 en Algoritmo 5). En el momento de la prueba, se considera los datos hasta la marca de tiempo actual y, por lo tanto, esta operación se ejecuta secuencialmente en línea. Una vez que se tiene los puntajes de anomalía para una marca de tiempo para cada dimensión s_i , se etiqueta la marca de tiempo como anómala si esta puntuación es mayor que un umbral. En **algoritmo 5**, se presenta la fase de prueba para TranAD presentado por Tuli et al. (2022).

Algorithm 5 The Tran AD testing algorithm

Require:

Trained Encoder E , Decoders D_1 and D_2

Test Dataset \widehat{W}

1: **for** ($t = 1$ to \widehat{T})

2: $O_1, O_2 \leftarrow D_1(E(\widehat{W}_t, \vec{0})), D_2(E(\widehat{W}_t, \vec{0}))$

3: $\widehat{O}_2 \leftarrow D_1(E(\widehat{W}_t, \|O_1 - W\|_2)), D_2(E(\widehat{W}_t, \|O_1 - W\|_2))$

4: $s = \frac{1}{2} \|O_1 - \widehat{W}\|_2 + \frac{1}{2} \|\widehat{O}_2 - \widehat{W}\|_2$

5.6.6 Entrenamiento del modelo

Por lo general, se recomienda que los conjuntos de datos de entrenamiento y validación utilizados para los modelos sin supervisión no contengan pocas anomalía. Esto se debe a que el modelo está diseñado para detectar anomalías en los datos, por lo que no aprenderá a detectarlas si los datos ya contienen anomalías. Para garantizar que el modelo aprenda a detectar anomalías, es mejor utilizar conjuntos de datos normales de funcionamiento sin anomalías para el entrenamiento y la validación.

El entrenamiento consta de varios pasos. En primer lugar, la clase de entrenamiento crea una instancia del modelo con todas las configuraciones (hiperparámetros). Luego, invoca el método de entrenamiento del modelo. Poco después de que el modelo se ha entrenado, la clase guarda un modelo entrenado en el sistema de archivos para su uso posterior. Durante el entrenamiento, el modelo se valida en cada época. El modelo registra los errores de entrenamiento y validación para su análisis posterior. Finalmente, después del entrenamiento el modelo vuelve a predecir el conjunto de validación para obtener el error de validación del modelo. Este error de validación se utiliza para determinar el umbral para clasificar las etiquetas. Aquí, se calcula y guarda en el sistema de archivos la media y desviación estándar de la validación para su uso posterior.

5.6.7 Procedimiento de evaluación del modelo

Una vez entrenado, el modelo se sometió a pruebas con datos no utilizados en el proceso de entrenamiento. Luego, el umbral se puede usar para generar las etiquetas de anomalía para los conjuntos de datos. Después de generar las puntuaciones de anomalías, las estadísticas esenciales se generan utilizando etiquetas de anomalías predichas utilizando las métricas de desempeño: precisión, recall y F1-Score.

6. Resultados

En este capítulo, se describen los diferentes experimentos llevados a cabo con un conjunto de datos del sistema SCADA de la CEV. Todos estos experimentos se llevaron a cabo utilizando el lenguaje de programación Python y utilizando algunas librerías adicionales para mejorar los resultados.

6.1 Hardware

Los experimentos fueron desarrollados en el entorno de Google Colab que en su versión gratuita ofrece una CPU Intel Xeon a 2.20 GHz, 13 GB de RAM, acelerador Tesla K80 y 12 GB de VRAM GDDR5.

6.2 Software

Todo lo descrito en esta sección fue realizado utilizando el lenguaje de programación Python. El código fuente de todo el proyecto está disponible bajo petición dirigida al Centro de Investigaciones Tecnológicas y Energéticas (CITE). Por ser el código demasiado largo no se incluyó en los anexos.

6.3 Procesamiento de datos

En esta investigación, se utilizó un conjunto de datos de la CEV, proporcionado por Unidad de Negocios GENSUR EP, que consistió en un archivo comprimido con subcarpetas organizadas por año, aerogenerador y día. En los datos en crudo se encontró diferentes desafíos que se detallan en la **Tabla 6**.

Tabla 6. Desafíos encontrados en los archivos del SCADA.

Descripción	N°
Archivos Procesados TXT	21 263
Archivos Procesados TMP	8 016
Archivos sin encabezado	3
Archivos con datos faltantes o con error de escritura	5
Total de archivos procesados	29 287

Para procesar los datos y generar un archivo CSV con la información de todos los aerogeneradores, se aplicó la metodología descrita en la **Figura 8**. Al realizar el análisis de los datos se encontraron 44 filas con valores nulos en todas las variables, dado que estas filas eran pocas y no afectaban significativamente al conjunto de datos, se las eliminó. Además, se encontraron las columnas como “today_data_energy_yield”, “yesterday_data_energy_yield”, “two_day_before_data_energy_yield” y “three_days_before_data_energy_yield”, las cuales contenían 1 820 394 datos nulos, y debido a que son significativos, se decidió eliminar estas columnas. Los detalles del conjunto de datos procesado se presentan en la **Tabla 7**.

Tabla 7. Datos del sistema SCADA.

timestamp	wind_speed_avg	wind_speed_max	wind_speed_min	grid_active_power_avg	data_power_production_time	two_day_before_data_energy_yield	three_days_before_data_energy	WT
1/1/2014 0:00	12.61	16.65	6.71	1293.90	34972.54	6151.0	4678.0	1
1/1/2014 0:10	13.10	17.17	8.51	1405.68	36972.04	4844.0	6151.0	1
.....									
.....									
31/10/2021 23:40	10.02	16.82	6.78	1283.90	44976.34	4844.0	6151.0	11
31/10/2021 23:50	13.10	17.20	8.30	1705.67	54972.04	4847.0	7151.0	11

El dataset final contiene 67 variables, que representan 4 123 012 datos para los 11 aerogeneradores, desde el 1 de enero de 2014 hasta el 31 de octubre del 2021 en frecuencia de 10 minutos.

6.4 Integridad de los datos

En primer lugar, con base en la revisión de la literatura consultada, se eliminaron las variables que no aportan con información relevante para el estudio. Estas variables se encuentran descritas en la **Tabla 8**.

Tabla 8. Variables eliminadas que no serán parte de este proceso.

ID	Variables que no formaron parte de este proceso
54	data_energy_yield
55	data_consumed_energy_yield
56	data_Power_on_time
57	data_WTG_ok_time
58	data_error_time
59	data_environment_ok_time
60	data_environment_not_ok_time
61	data_service_time
62	data_grid_control_standstill_time
63	data_power_production_time
64	today_data_energy_yield
65	yesterday_data_energy_yield
66	two_day_before_data_energy_yield
67	three_days_before_data_energy_yield

6.5 Filtrado de información del conjunto de datos

Para el filtrado de datos, se procedió según lo indicado en el apartado **5.6.2**, que recomienda lo siguiente:

- Potencia debe ser mayor a cero
- Velocidad del viento mayor a 3m/s (según especificaciones técnicas de velocidad de arranque)
- Potencia activa menor a 1600 W para evitar mediciones no adecuadas.

- Velocidad del viento menor a 26 m/s, que es la velocidad máxima de corte.
- Velocidad del generador mayor a cero.
- Temperatura máxima del IGBT menor a 120 °C, que es el doble de la media.
- Temperatura mínima de 0°C
- Temperatura máxima del CHOPPER IGBT menor a 60 °C, que es el doble de la media.
- Temperatura máxima del IGBT menor a 120 °C, que es el doble de la media.
- Modo de operación 5, que es el modo en el que el aerogenerador se encuentra en funcionamiento normal.

Se tomaron en cuenta estos aspectos con el objetivo de limitar los datos de entrenamiento para cuando la turbina está en estado de funcionamiento y no en un estado de reposo.

6.6 Etiquetado de datos

Según el registro de mantenimiento, se identificaron y marcaron los incidentes de fallas y alarmas del módulo IGBT en los datos SCADA de los aerogeneradores, siguiendo el proceso especificado en la **Figura 9**. Para el filtrado de fallas y alarmas en el archivo de alarmas del SCADA y el registro de O&M, se utilizaron los siguientes criterios:

- Elección de los códigos de alarmas estrechamente relacionadas con el módulo IGBT y convertidor.
- Recomendación de literatura científica y de expertos en el área.

Los nombres y códigos de alarma utilizados en este estudio se detallan en la **Tabla 9**.

Tabla 9. Filtrado de fallas y código de alarmas.

Datos	Componente	Nombre de falla o código de alarma
Archivo de operación y mantenimiento	IGBT	MÓDULO IGBT, TIPO GOLDWIND 1500KW-G-05 TECHWIN
Archivo de alarmas SCADA	IGBT	435,442,496,503,451,506,447,453,448,523,420,596,452,484,483,509,446,454,497,444,508,433,531,449,422,424,458,512,482,481,486,485,504,505,432,597,507,510,519,470

En el **Anexo 2**, se detallan los códigos de las alarmas del IGBT y convertidor. Solo se aborda este componente, ya que ha sido el más reemplazado, mientras que los demás siguen siendo desconocidos o no existen datos suficientes, debido a que los aerogeneradores son relativamente nuevos. En la **Tabla 10** se muestra el número de alarmas, fallas y conjunto de datos correspondientes a cada aerogenerador.

Tabla 10. Datos SCADA, alarmas y fallas del módulo IGBT.

WT	Data	IGBT Alarms	IGBT Faults
1	321042	352	3
2	317323	641	3
3	327283	274	3
4	322918	292	3
5	324174	417	2
6	320076	191	1
7	333590	229	3
8	338179	243	1
9	327177	240	1
10	318669	236	3
11	231750	180	2
Total	3482181	3295	25

En la **Tabla 11** se expone una muestra de las variables de estudio, y en las últimas 3 columnas se muestra el aerogenerador asociado, el estado de alarmas y fallas para el módulo IGBT, correspondiendo el valor de 1 a un estado no saludable (falla), y 0 a saludable (normal).

Tabla 11. Etiquetado de fallas de los datos SCADA.

timestamp	wind_speed_avg	wind_speed_max	wind_speed_min	grid_active_power_avg	three_days_before_data_energy	WTG	Alarm_IGBT	Fault_IGBT
1/1/2014 0:00	12.61	16.65	6.71	1293.90	4678.0	1	0	0
1/1/2014 0:10	13.10	17.17	8.51	1405.68	6151.0	1	1	1
31/12/2021 23:40	10.02	16.82	6.78	1283.90	6151.0	11	1	0
31/12/2021 23:50	13.10	17.20	8.30	1705.67	7151.0	11	1	1

6.7 Análisis estadístico

Se realizó un análisis estadístico de los datos del sistema SCADA, para comprender donde se producen las fallas y las alarmas. De aquí en adelante se usará el ID de las variables para referirse a ellas.

6.7.1 Descripción de las variables de estudio

Se utilizó la librería de Pandas para extraer las principales medidas de tendencia central de las principales variables del sistema SCADA. En la **Tabla 12** se presentan estos valores.

Tabla 12. Descripción de valores de tendencia central.

ID Feature	Mean	Std	min	max
1	10.87	4.06	1.47	21.98
4	923.56	555.46	0.01	1578.78
7	23.90	4.79	12.60	41.30
8	-0.02	0.41	-9.97	9.99
11	16.30	3.05	0.26	19.04
14	13.39	1.68	8.11	23.64
17	2.63	5.12	-8.36	85.89

ID Feature	Mean	Std	min	max
20	-27.61	253.92	-776.30	726.60
21	360.44	3.78	343.26	381.27
24	881.91	518.20	14.31	1567.87
28	50.93	19.09	14.75	82.40
29	19.41	1.80	13.00	31.40
30	13.66	1.87	9.00	26.60
31	62.37	18.52	31.81	94.86
32	26.34	7.01	11.90	49.70
33	71.94	17.00	15.60	104.10
34	35.31	8.43	11.60	52.50
35	32.28	0.38	31.23	36.85
36	50.18	6.02	14.20	72.20
37	53.29	14.53	31.52	82.44
38	24.25	10.98	10.20	51.60
41	24.20	1.98	15.80	30.90
45	21.84	2.06	14.00	29.00
47	20.12	3.36	14.00	29.90
50	21.84	2.06	14.00	29.00

En el **Anexo 1** se encuentra el ID y su descripción de cada variable.

6.7.2 Curva de potencia y alarmas

En la **Figura 13** se muestra los datos normales (puntos azules) de potencia y velocidad, y se observan las alarmas del módulo IGBT en color rojo. Además, se puede apreciar que la mayoría de las alarmas ocurren cuando el aerogenerador alcanza su potencia nominal.

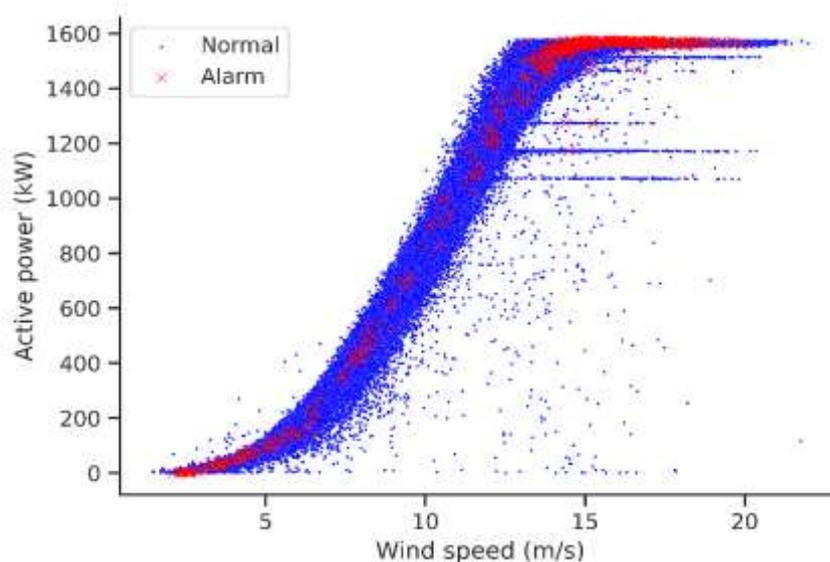


Figura 13. Curva de potencia WT1.

A continuación, en la **Figura 14** se presenta el mismo patrón para todos los aerogeneradores en cuanto a la tendencia de las alarmas en relación a la curva de potencia.

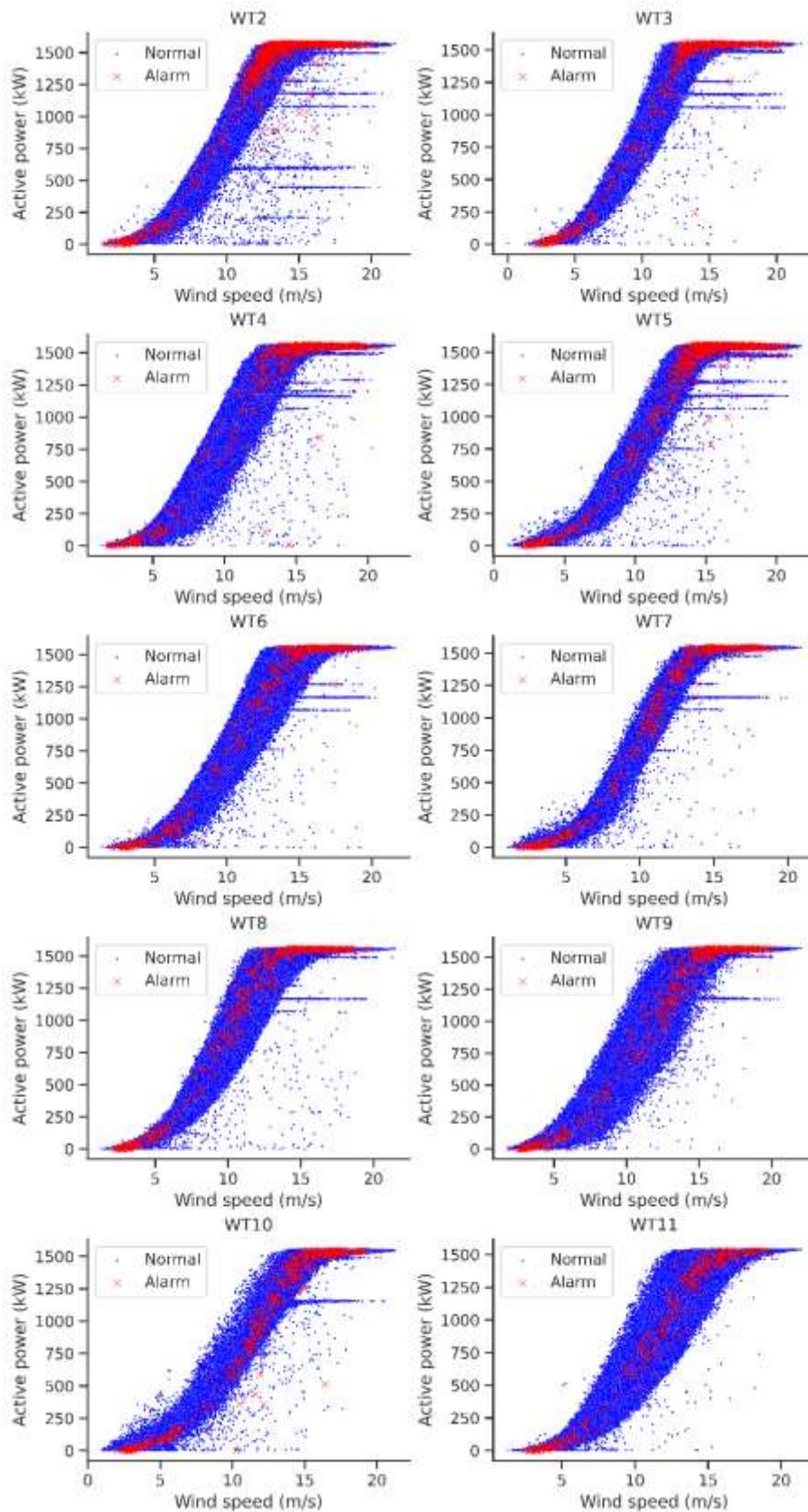


Figura 14. Curva de potencia para todos los aerogeneradores.

6.7.3 Curva de Potencia vs Temperatura IGBT

En la **Figura 15** se observa claramente que cuando el aerogenerador 1 se encuentra operando a su potencia nominal, existe una concentración de datos de alarma relacionados con la temperatura del módulo IGBT. Este patrón de comportamiento se observa en todos los aerogeneradores como se puede observar el **Anexo 8**.

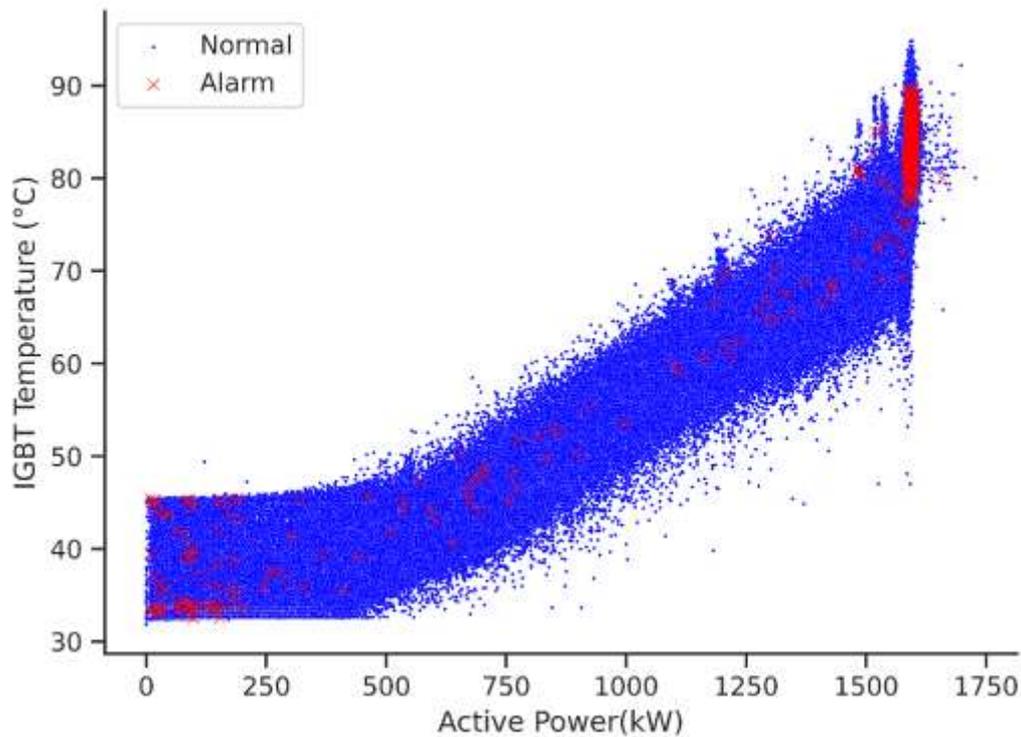


Figura 15. Curva de Potencia vs Temperatura IGBT.

6.7.4 Relación de las variables a través del tiempo

En la **Figura 16** se presenta la relación entre cinco variables en serie temporal, que están correlacionadas con las fallas en el módulo IGBT. Estas gráficas muestran los rangos en los que se encuentran estos valores y además se indica con una X de color rojo cuándo ocurrió la falla. Al observar el comportamiento de las variables, no es posible determinar a simple vista la causa de la falla, ya que no se observa ninguna diferencia significativa entre los datos anteriores y los datos en el momento de la falla.

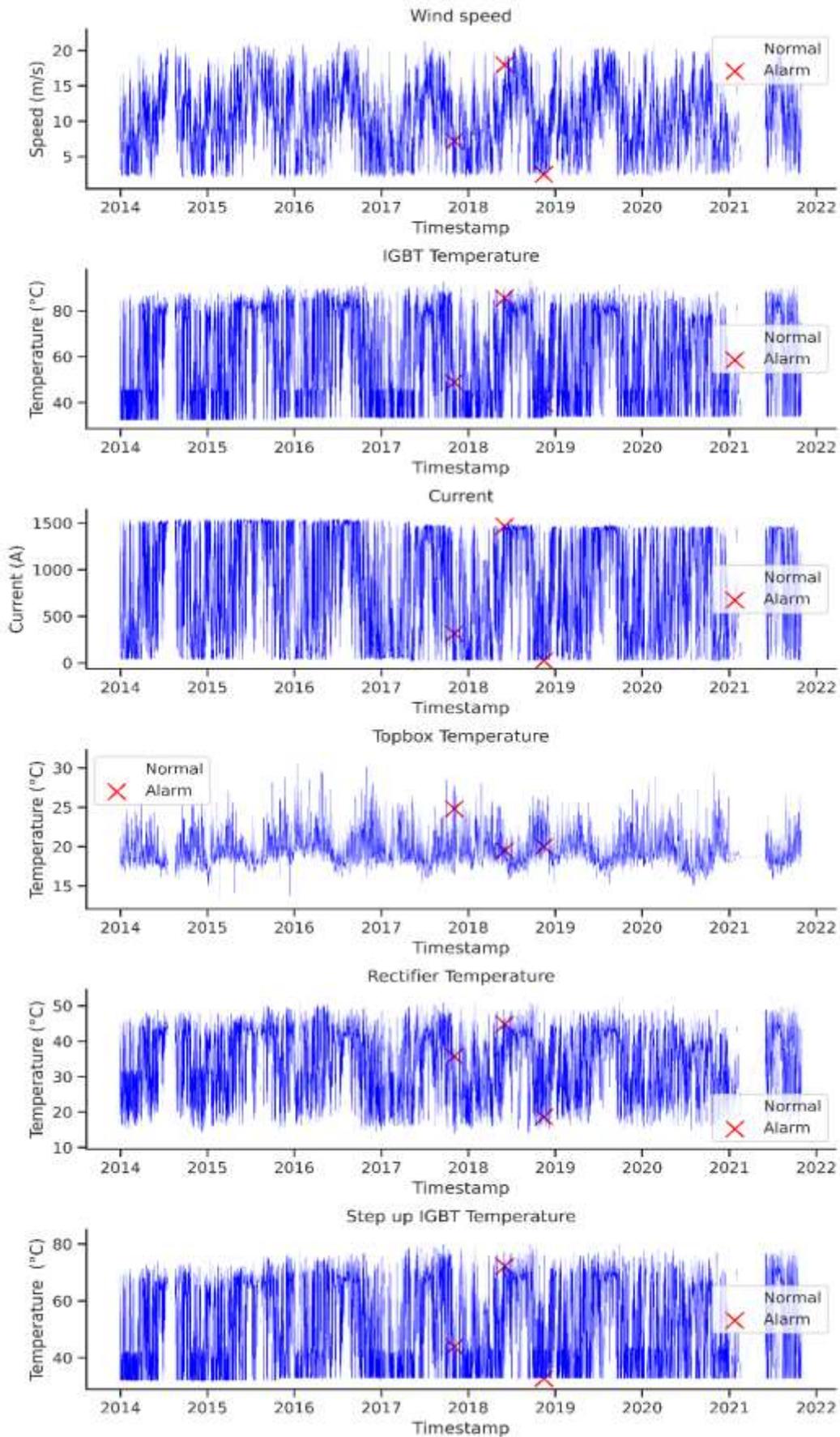


Figura 16. Serie temporal de variables relacionadas con la falla del módulo IGBT en WT1.

6.7.5 Gráfico de control

La **Figura 17** muestra el gráfico de control correspondiente a la variable 'igbt_temperature_max' en la WT1, en donde se observan las fallas con una X de color rojo y el comportamiento de esta variable con un umbral establecido en la media de los datos más 1,5 desviaciones estándar.

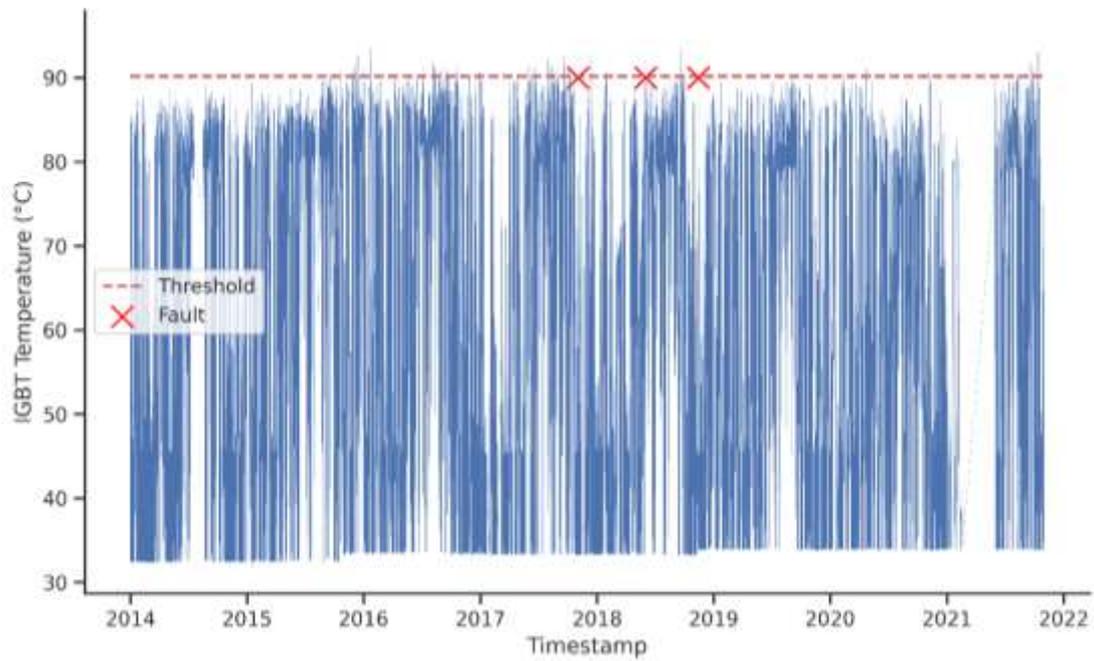


Figura 17. Serie temporal de la temperatura del IGBT

6.7.6 Gráfico de caja y bigote

En este gráfico se observó el rango de los datos para cada variable, los valores medianos, los cuartiles y cualquier valor atípico que pueda estar presente. Este diagrama es útil para visualizar rápidamente la distribución del conjunto de datos y permitió identificar valores atípicos y posibles problemas con los datos. Además, se lo usó para comparar entre conjuntos de datos de todos los aerogeneradores donde se buscó similitudes y diferencias entre ellos.

En la **Figura 18** se observa el diagrama de caja y bigotes para el WT1, y se observa que existen muchos valores atípicos en la variable de potencia reactiva, ángulo de las palas y aceleración del nacelle, esto puede ser debido a la inestabilidad de los aerogeneradores en condiciones de operación y la variación del viento. Además, estos valores atípicos se pueden presentar debido a que los aerogeneradores no generan siempre la misma cantidad de energía, sino que varían dependiendo del viento, esto puede provocar una mayor variabilidad en la potencia reactiva, lo que se refleja en la presencia de valores atípicos en el diagrama de caja y bigotes.

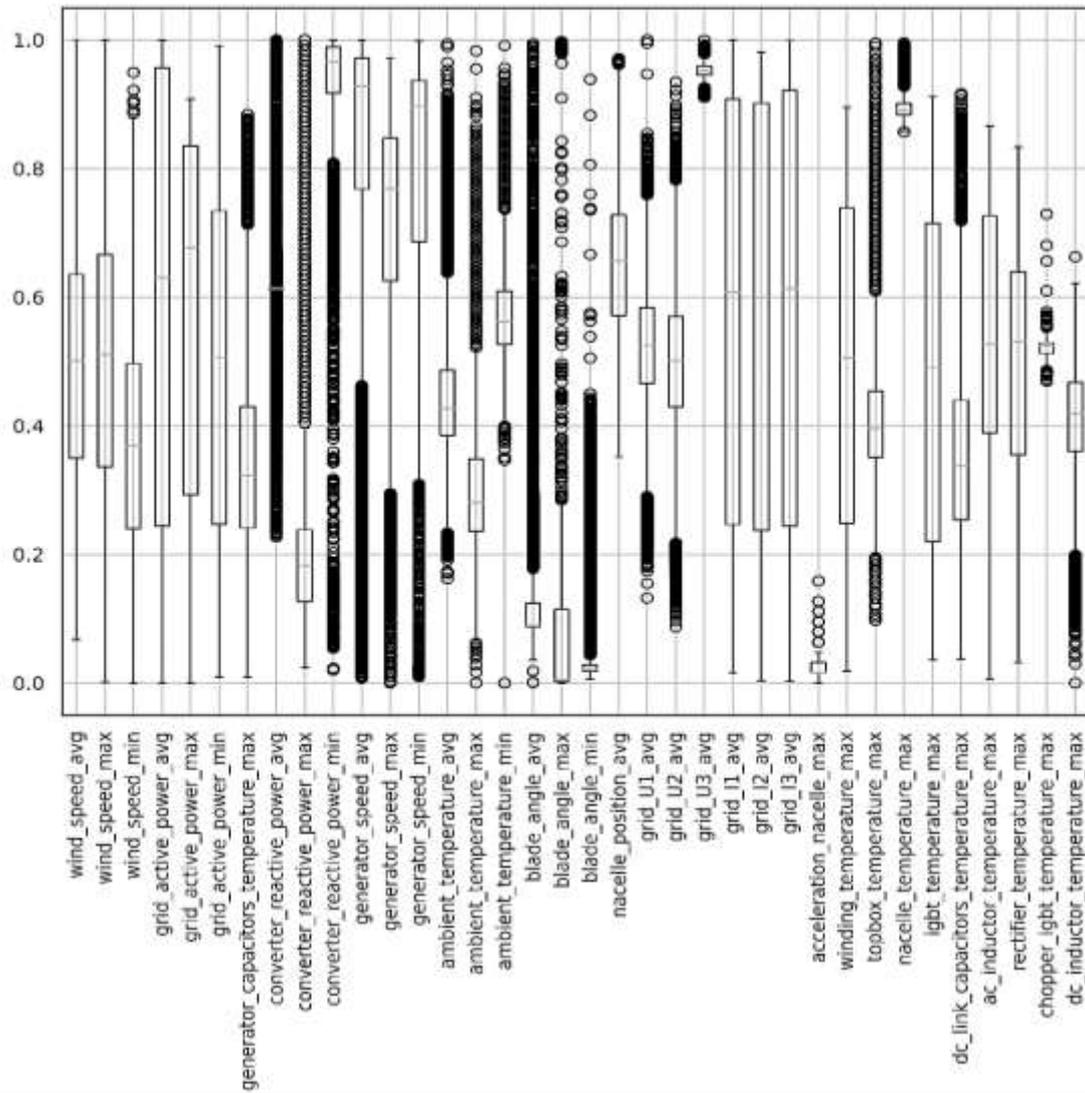


Figura 18. Diagrama de caja y bigotes de las variables del sistema SCADA.

6.8 Selección de características

6.8.1 Método de eliminación recursivo de características (RFE).

Al aplicar el método RFE, se combinaron varias variables basadas en su relación con el componente que presentaba fallas. Los resultados mejorados permitieron seleccionar las variables del sistema SCADA, presentadas en la **Tabla 13**.

6.8.2 Método estadístico de correlación (C)

Esta matriz es útil para determinar la correlación entre diferentes parámetros tales como temperatura, potencia, velocidad del viento, y otras variables importantes. Esta herramienta permite explorar la relación entre las variables de estudio, la temperatura del IGBT (*igbt_temperature_max*), y diferentes variables como la velocidad del viento (*wind_speed_avg*), la corriente de la red (*grid_I1_avg*), la temperatura de la caja de control, entre otras. Esta información proporciona una visión rápida de las relaciones entre estas

variables y ayuda a comprender de mejor manera, cómo interactúan entre sí para poder utilizar las variables adecuadas en el modelo de predicción fallas en los aerogeneradores.

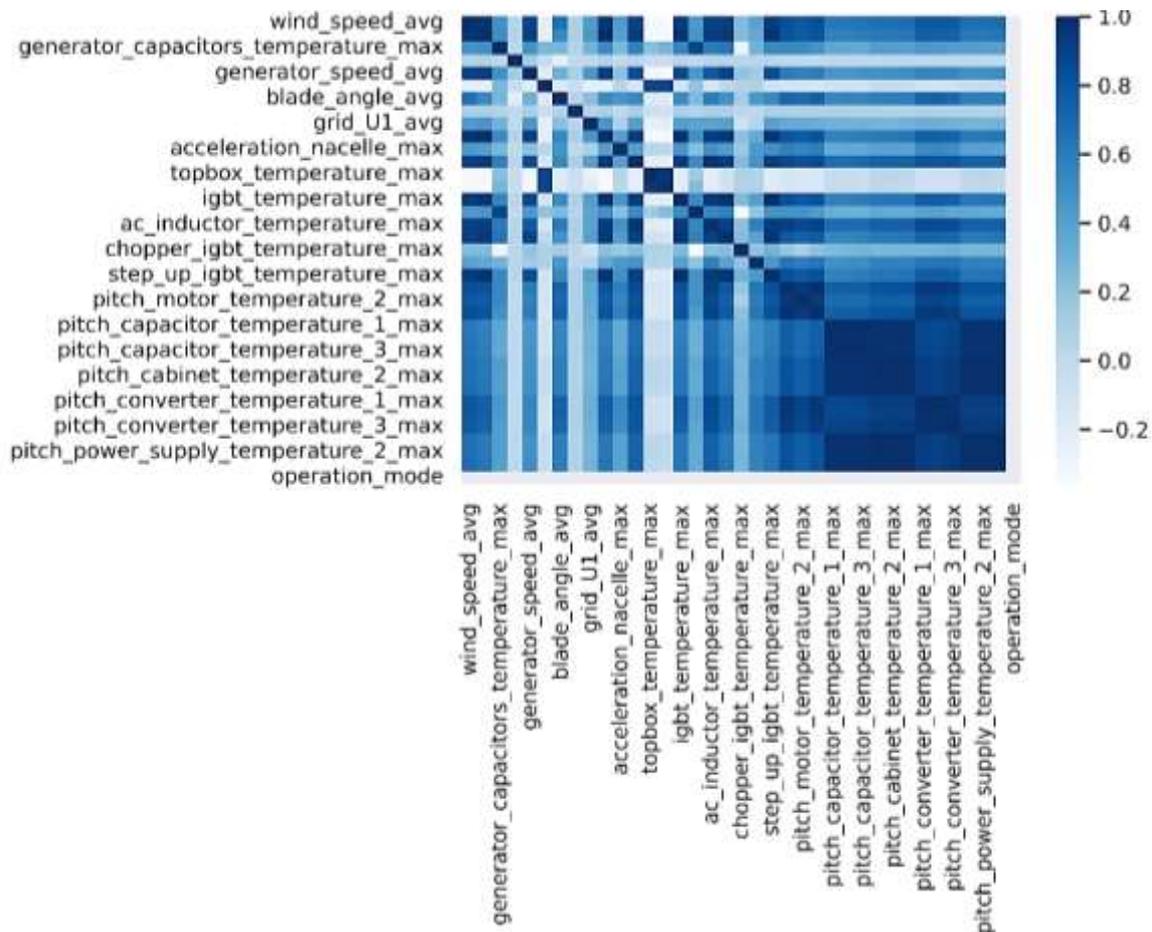


Figura 19. Matriz de correlación de las variables del sistema SCADA.

Al analizar la matriz de correlación, se identificaron las variables relacionadas con la falla en el IGBT. La variable clave para el análisis fue la temperatura máxima del IGBT (*igbt_temperature_max*). Se examinaron las variables con un índice de correlación de Pearson mayor a 0,75 (Ver **Figura 19**), resultando en las siguientes variables con ID: 1, 2, 3, 4, 5, 6, 8, 9, 10, 14, 15, 16, 21, 22, 23, 24, 25, 26, 29, 31, 33, 34, 36 y 37.

6.8.3 Método estadístico de prueba de hipótesis (PV)

El método de Prueba de Hipótesis se utilizó con las variables SCADA y las alarmas IGBT etiquetadas para identificar las variables más relevantes con valores p inferiores a 0,05. El código en Python para la selección de variables se encuentra en el libro de Ozdemir & Susarla (2018). Las variables seleccionadas con este método son las siguientes con ID: 17, 18, 27, 19, 21, 6, 33, 25, 26, 24, 4, 31, 16, 34, 5, 28, 29 y 31.

6.8.4 Métricas de selección basado en árboles (AD)

Para la selección de características con el método basado en árboles, se utilizaron las variables SCADA y las alarmas del IGBT. Debido a que las fallas en el IGBT ocurren entre 2 a 4 veces por cada aerogenerador, estas alarmas fueron consideradas clave para la selección. Los ID de las variables seleccionadas son: 1, 2, 3, 4, 5, 6, 7, 8, 9, 15, 16, 18, 19, 20, 21, 22, 24, 25, 28, 30, 31, 33 y 34. El código para este método se encuentra en el libro de Ozdemir y Susarla (2018).

6.8.5 Modelos lineales para selección de características (ML)

Se utilizó la regresión logística, siguiendo el enfoque de Ozdemir y Susarla (2018), para seleccionar variables a partir de los datos operativos del SCADA y la alarmas del IGBT. Como resultado, se identificaron las siguientes variables con ID: 1, 2, 3, 4, 14, 15, 17, 18, 19, 21, 24, 25, 26, 28, 29, 31, 33 y 36.

Una vez realizado el proceso de selección de variables, se presenta un resumen en la siguiente tabla.

Tabla 13. Resumen de las variables seleccionadas por diferentes métodos.

FS	N° Features	ID Feature
RFE1	1	31
RFE12	12	2,4,8,15,21,24,29,31,33,34,36,37
RFE24	24	1,2,3,4,5,6,8,9,10,14,15,16,21,22,23,24,25,26,29,31,33,34,36,37
C6	6	2,5,24,29,31,37
PV18	18	17,18,27,19,21,6,33,25,26,24,4,31,16,34,5,28,29,31
AD23	23	1,2,3,4,5,6,7,8,9,15,16,18,19,20,21,22,24,25,28,30,31,33,34
LM18	18	1,2,3,4,14,15,17,18,19,21,24,25,26,28,29,31,33,36

De ahora en adelante, FS se utilizará para referirse al método de selección que se utilizará como entrada al modelo. Cada método viene acompañado de un valor; por ejemplo, C6, donde C corresponde al método de selección por correlación y 6 es el número de variables resultantes de este método.

6.8.6 Escalado de características

La técnica de escalado de variables es importante en el proceso de procesamiento de datos para asegurar que las variables numéricas sean comparables entre sí, ya que diferentes variables pueden tener diferentes unidades de medida o rango de valores, lo que puede afectar negativamente a algunos algoritmos de aprendizaje automático. Se utilizó la librería de Sklearn y el módulo MinMaxScaler para escalar las variables, donde se ajustó el rango de valores de cada variable a un rango específico entre 0 y 1.

6.8.7 División de los datos

Como se utilizó modelos de aprendizaje profundo sin supervisión para analizar los datos, se requería que el conjunto de entrenamiento y validación contuvieran un mínimo de anomalías para poder entrenar y validar los modelos con datos normales. Por otro lado, se utilizaron datos conocidos con fallas en el módulo IGBT para realizar las pruebas. La presentación de los resultados de la división de datos se encuentra representada en la **Tabla 14**, donde se estableció como regla general entrenar los modelos con datos obtenidos 1 año antes de la ocurrencia de la primera falla en cada aerogenerador. Finalmente, se utilizó el conjunto de entrenamiento para entrenar el modelo, el conjunto de validación para evaluar su desempeño y ajustar los hiperparámetros, y el conjunto de pruebas para evaluar el desempeño del modelo final.

Tabla 14. División de los datos para los 11 aerogeneradores.

WT	Split	Start Train	Train	Val	Start test	Test
1	0.57	1/1/2014 0:00	138049	0.30	11/3/2017 19:20	182993
2	0.52	1/1/2014 0:00	152316	0.30	4/7/2017 21:40	165007
3	0.40	1/1/2014 0:00	196370	0.30	22/5/2018 10:50	130913
4	0.57	1/1/2014 0:00	138855	0.30	11/3/2017 8:20	184063
5	0.34	1/1/2014 0:00	213955	0.30	30/10/2018 22:50	110219
6	0.50	1/1/2014 0:00	160038	0.30	2017-08-23 03:50:00	160038
7	0.48	1/1/2014 0:00	173467	0.30	10/10/2017 16:20	160123
8	0.54	1/1/2014 0:00	152181	0.30	6/4/2017 19:20:00	185998
9	0.40	1/1/2014 0:00	196307	0.30	17/5/2018 21:10	130870
10	0.40	1/1/2014 0:00	191202	0.30	13/5/2018 0:10	127467
11	0.30	1/1/2014 0:00	185400	0.30	5/3/2020 06:30	46350

6.9 Aplicación del primer modelo de detección de anomalías (Anomaly Transformer)

6.9.1 Tamaño de ventana deslizante

Es fundamental tener en cuenta el tamaño de la ventana que se utiliza para el análisis de datos. Este parámetro determina la profundidad temporal del modelo y es crítico para garantizar la eficacia del análisis. Un tamaño de ventana demasiado pequeño puede resultar en la captura de patrones locales de corta duración, mientras que un tamaño demasiado grande puede dificultar la detección de patrones temporales específicos, debido al procesamiento de ventanas de tiempo excesivamente amplias. Además, es importante tener en cuenta que el tamaño de la ventana puede afectar la precisión de los resultados del análisis y, por lo tanto, es importante elegir un tamaño adecuado que tenga en cuenta los objetivos del análisis y los patrones de datos esperados.

6.9.2 Ventana deslizante para detección de fallas (primera etapa)

Cómo los modelos son capaces de detectar anomalías se utilizó el método de ventana deslizante para obtener una alarma antes de que se produzca un fallo real. Este enfoque implica que, si se detecta una anomalía, se registra hasta que se alcanza un umbral preestablecido. Si este umbral se supera en un determinado número de datos, se considera que existe una probabilidad de falla en el módulo IGBT en una ventana tiempo $tf = 7 \text{ meses}$. Es decir, si en N datos se supera un umbral de $thresh$ anomalías, se activa una alarma de posible falla en el módulo IGBT. Además, se establece una segunda ventana en la que, si en $N2$ datos ($N \gg N2$) se supera un umbral $thresh2$ anomalías, también es un indicativo de alarma de falla del módulo IGBT. Además de esto, para la ventana de tiempo 1 y 2 se usa otra ventana a y $a2$, respectivamente como rango de que, si se evalúa otra alarma dentro de este, se asume que pertenecen a la misma falla. Cabe recalcar que los valores de los parámetros son diferentes para cada método de selección de características y cada WT evaluado. Estos cambios pueden deberse a diferentes características seleccionadas, así como de la naturaleza inherentemente variable de las condiciones del viento.

6.9.3 Evaluación de detección de fallas

En la evaluación de eventos discretos del desempeño, un Verdadero Positivo (TP) corresponde a una advertencia generada por el modelo, asociado con un componente que presenta una falla perteneciente al módulo IGBT registrada en el archivo de O&M del WT en estudio, en una ventana de tiempo $tf = 7 \text{ meses}$. Por otro lado, se consideran Falsos Positivos (FP) a las advertencias del modelo que no están seguidas por una falla relacionada con el componente involucrado dentro de la ventana tf , Aquellas fallas registradas en O&M que no son predichas por una advertencia del modelo dentro de la ventana de tiempo de referencia se consideran Falsos Negativos (FN). El rendimiento del modelo se cuantificó a través de métricas de clasificación ampliamente utilizadas en el campo de aprendizaje automático, como precisión, Recall y el F1 Score, sus formulaciones matemáticas se presentan en las **ecuaciones (4), (5) y (6)** respectivamente.

6.9.4 Detección de fallas en WT1 mediante Anomaly Transformer

Tenga en cuenta que, aunque se presenta la relación de las fallas del IGBT con la variable de temperatura del IGBT, el modelo de Anomaly Transformer considera todas las variables de los métodos FS y su relación en el tiempo.

6.9.4.1 Anomaly Transformer utilizando el método RFE1 en WT1

En este caso se eligió la variable que mide la temperatura del módulo IGBT como la variable principal para analizar. Los parámetros de entrada utilizados en el modelo se resumen en la **Tabla 15**. Es importante mencionar que, al detectar anomalías en una sola variable, se está limitando el alcance del análisis y es posible que se pierdan patrones de fallas relacionadas con otras variables.

En la **Figura 20** se observa que el modelo advirtió de una falla en un avance promedio aproximado de 3 meses antes de que ocurra alguna de las tres fallas del módulo IGBT. Sin embargo, también se pueden ver una advertencia detectada donde no hubo falla a finales del 2020, dando un FP. Esto no significa necesariamente que el modelo haya fallado, sino que puede ser que otro componente este afectando para que se de este tipo de alarmas. Las fallas ocurridas el 02 de noviembre del 2017, fueron advertidas con 3 meses de anticipación aproximadamente mientras que para la falla ocurrida el 02 de junio de 2018, el modelo logró advertir 1 mes antes de la falla aproximadamente. A partir de aquí, en los resultados de este modelo se va a representar el umbral mediante una línea discontinua roja, y las anomalías serán sombreadas de color rojo cuando correspondan a una advertencia del modelo de posible existencia de falla en el módulo IGBT en los próximos 7 meses subsiguientes. A continuación, se denominará AT al modelo Anomaly Transformer.

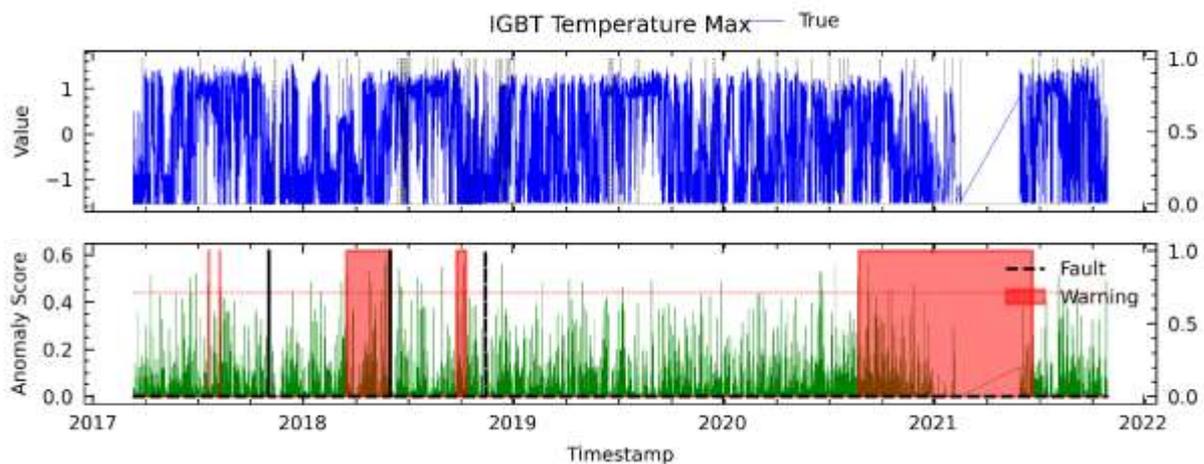


Figura 20. Resultados para RFE1 en el análisis de temperatura del módulo IGBT mediante el modelo AT para WT1.

6.9.4.2 Anomaly Transformer utilizando el método RFE12 en WT1

En el apartado anterior, se realizó un análisis de una sola variable enfocado en la temperatura del módulo IGBT. En este apartado, se llevó a cabo un análisis multivariado donde se van a incluir las variables seleccionadas de RFE12. Los parámetros de entrada utilizados en

el modelo se resumen en la **Tabla 15**. Es importante mencionar que, al analizar varias variables simultáneamente, se tiene una visión más completa del sistema y se pueden detectar patrones de fallas relacionadas con varias variables.

En la **Figura 21** se puede apreciar que el modelo logró detectar la primera falla aproximadamente 4 meses antes de su ocurrencia que fue el 02 de noviembre de 2017. Del mismo modo, el modelo advirtió con anticipación sobre la segunda y tercera falla, con aproximadamente 3 y 1 meses respectivamente. Sin embargo, a partir de ese momento, se volvió a detectar una posible falla a finales de 2019 e inicios de 2020, pero según los informes de O&M no existió ninguna falla.

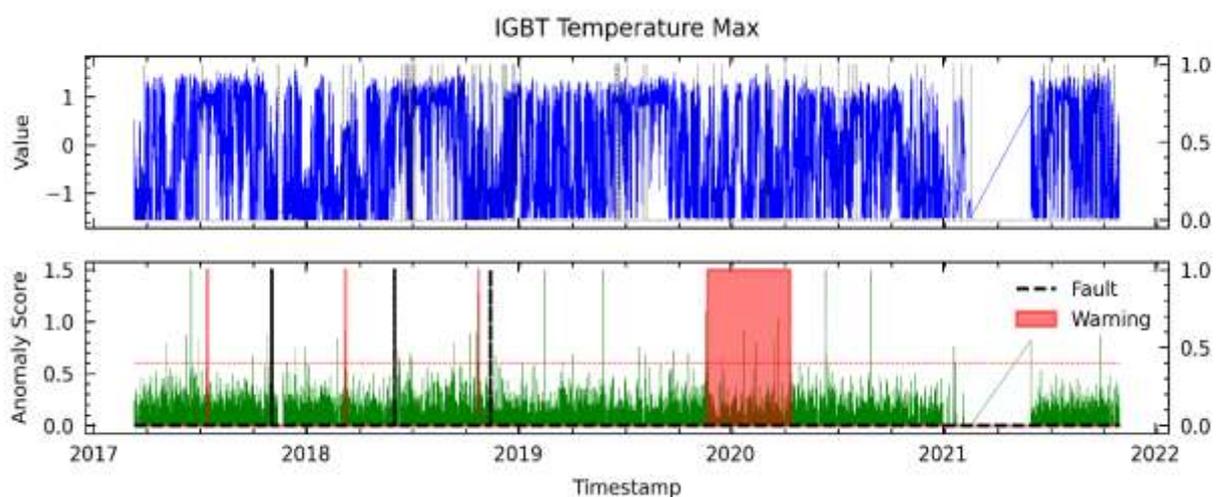


Figura 21. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.

6.9.4.3 Anomaly Transformer utilizando el método RF24 en WT1

En la **Figura 22** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que el modelo no logró detectar la primera falla, pero sí advirtió con anticipación sobre la segunda y tercera falla, con aproximadamente 2 y 1 meses de anticipación respectivamente. Sin embargo, después de ese momento, el modelo detectó advertencias sobre posibles fallas a finales de 2019, finales de 2020 e incluso los primeros meses de 2021, pero según los informes de O&M no se presentaron fallas.

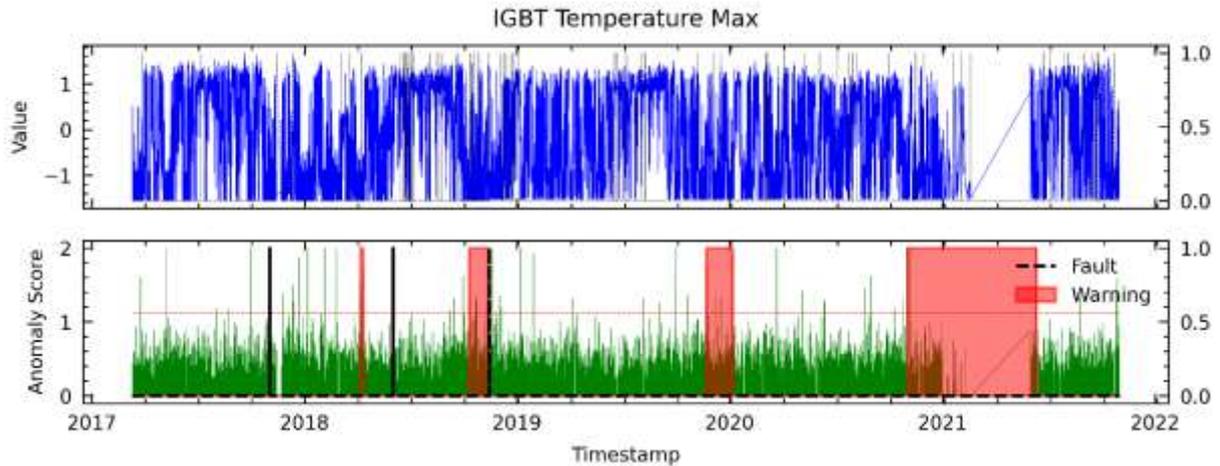


Figura 22. Resultados del modelo para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.

6.9.4.4 Anomaly Transformer utilizando el método C6 en WT1

En la **Figura 23** se puede observar que el modelo logró detectar las tres fallas del módulo IGBT con un promedio aproximado de seis meses de anticipación, pero también reportó dos falsos positivos (FPs).

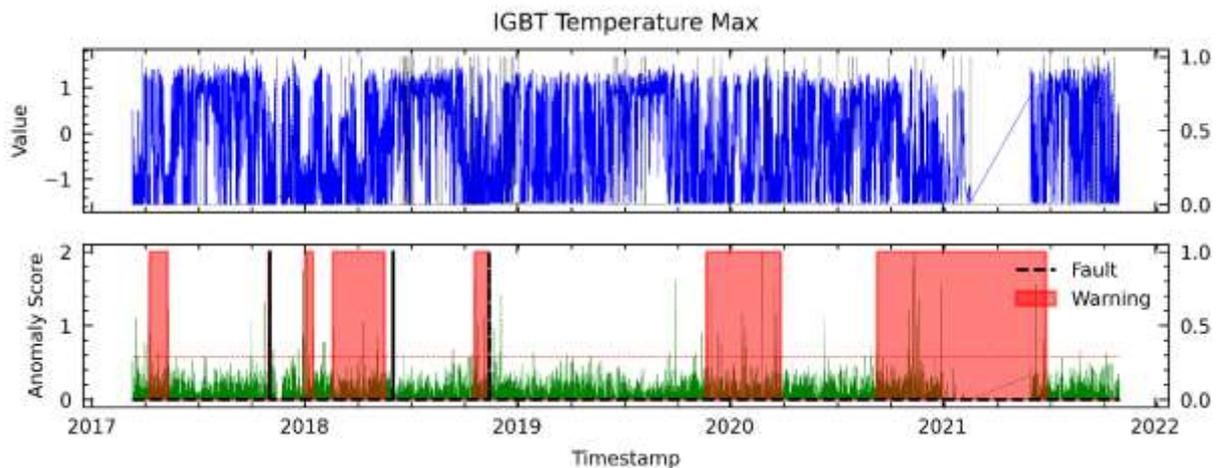


Figura 23. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.

6.9.4.5 Anomaly Transformer utilizando el método PV18 en WT1

En la **Figura 24** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra las advertencias de falla del módulo IGBT con una anticipación promedio aproximada de 3 meses antes de la ocurrencia de la falla. Además, se registró un caso de falsa alarma (FP), que es similar a los resultados obtenidos con el método RFE12.

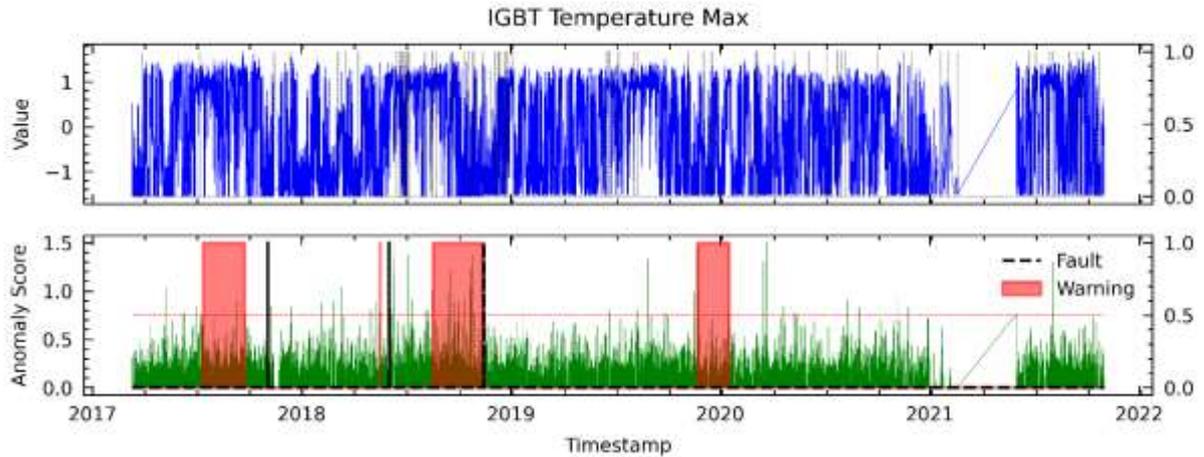


Figura 24. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.

6.9.4.6 Anomaly Transformer utilizando el método AD23 en WT1

En la **Figura 25** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados no son satisfactorios, ya que, aunque es capaz de detectar las tres fallas en el módulo IGBT con hasta tres meses de anticipación aproximadamente, también se han producido tres falsas alarmas.

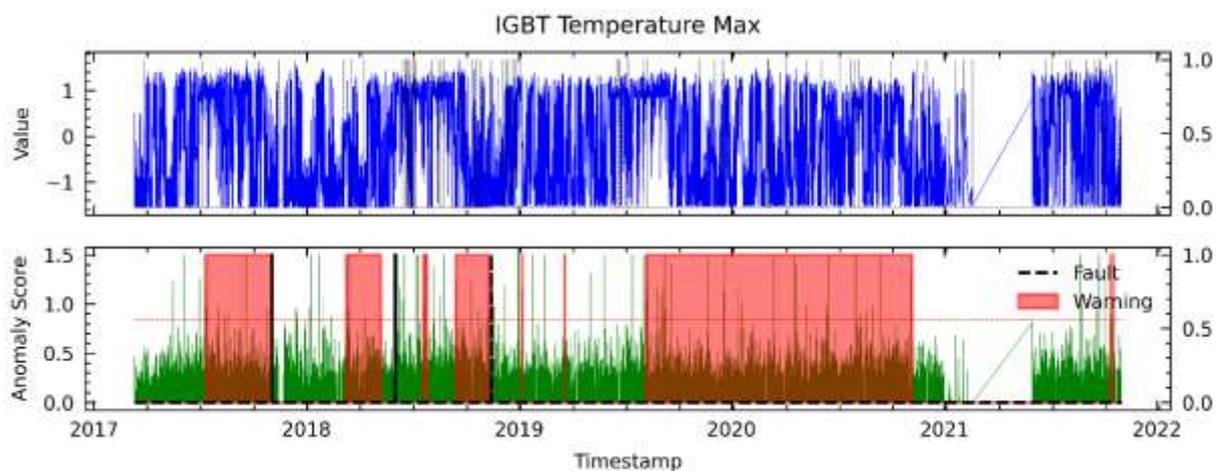


Figura 25. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.

6.9.4.7 Anomaly Transformer utilizando el método LM18 en WT1

En la **Figura 26** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que se lograron detectar las fallas en el módulo IGBT con hasta 3 meses de anticipación aproximadamente, pero también se presentaron 2 FPs, un resultado similar al obtenido con las variables seleccionadas por C6.

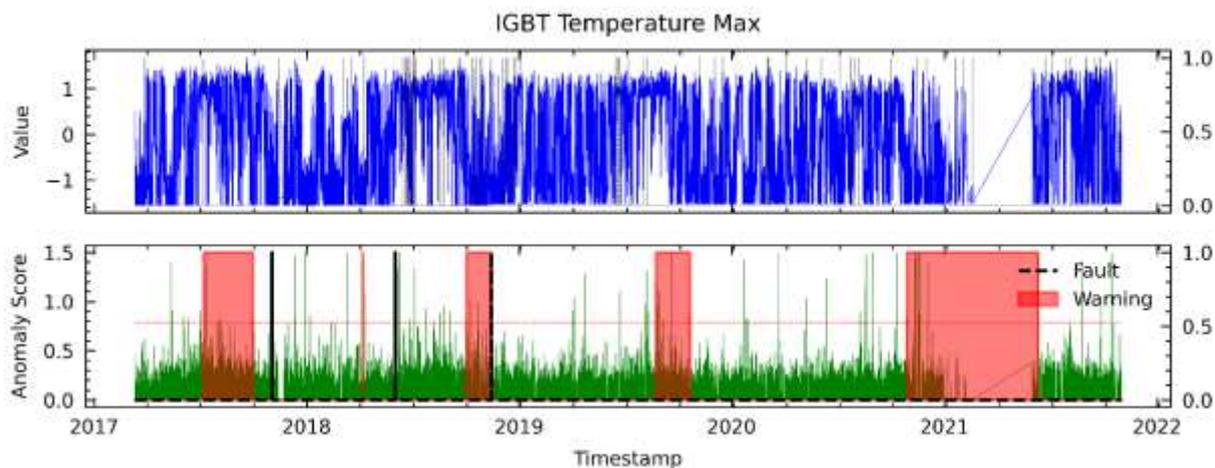


Figura 26. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT1.

6.9.4.8 Resultados de Anomaly Transformer para WT1

Los resultados para WT1 que se presentan en la **Tabla 15**, se obtuvieron con los siguientes hiper-parámetros: tamaño de ventana de 45, paso de 1, 10 épocas de entrenamiento y una tasa de aprendizaje de $1e^{-6}$. Estos hiper-parámetros descritos fueron utilizados para todos los WTs.

Tabla 15. Resultados de rendimiento del modelo AT para WT1.

WT	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	R	P	F1
	RFE1	0,015	12960	4	1000	2000	3	2000	3	1	0	0,75	1,00	0,86
	RFE12	0,100	10000	4	1000	2000	3	2000	3	1	0	0,75	1,00	0,86
	RFE24	0,100	12960	4	1000	2000	3	2000	2	2	1	0,50	0,67	0,57
1	C6	0,060	10000	4	1000	2000	3	2000	3	2	0	0,60	1,00	0,75
	PV18	0,135	12960	5	1000	2000	3	2000	3	1	0	0,75	1,00	0,86
	AD23	0,135	12960	5	1000	2000	3	2000	3	3	0	0,50	1,00	0,67
	LM18	0,100	12960	5	1000	2000	3	2000	3	2	0	0,60	1,00	0,75

Nota: El mejor rendimiento se presenta en negrita.

Donde, r es el umbral establecido en el percentil más bajo de Anomaly Score, N y $N2$ corresponden a la primera y segunda ventana de datos usada para la densidad de anomalías, $thresh$ y $thresh2$ son los umbrales de densidad que deben superarse para mostrar una advertencia de fallo en el módulo IGBT, a y $a2$ son los rangos de datos utilizados para determinar si se encuentra otra advertencia corresponde a la misma falla.

6.9.5 Detección de fallas en WT2 mediante Anomaly Transformer

A continuación, se presentan los resultados con la mejor puntuación de F1 Score con el objetivo de presentarle al lector los mejores resultados de los métodos desarrollados en la investigación. Para más detalles puede consultar el **Anexo 3**.

6.9.5.1 Anomaly Transformer utilizando el método RFE 12 en WT2

En la **Figura 27** se presenta el resultado para este método donde se puede apreciar que el modelo logró detectar la primera falla 3 meses antes de su ocurrencia aproximadamente, que fue el 10 de junio del 2018. Del mismo modo, el modelo advirtió con anticipación sobre la segunda falla, con 7 meses de anticipación aproximadamente. Sin embargo, a partir de ese momento, se volvió a detectar advertencias de posibles fallas a finales de 2021, pero no se tiene información de lo que ocurrió después, por lo que para esta evaluación se lo toma como FP.

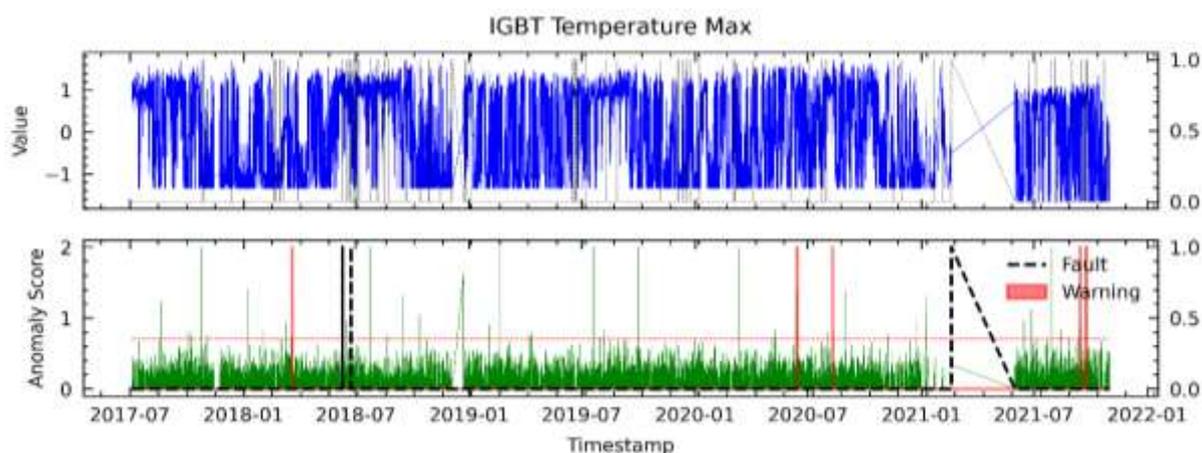


Figura 27. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.

6.9.5.2 Resultados de Anomaly Transformer para WT2

Los resultados para la WT2 se presentan en la **Tabla 16**.

Tabla 16. Resultados de rendimiento del modelo AT para WT2.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,040	12960	5	1000	2000	3	2000	2	2	0	0,50	1,00	0,67
	RFE12	0,115	12960	5	1000	2000	3	2000	2	1	0	0,67	1,00	0,80
	RFE24	0,135	12960	5	1000	2000	3	2000	2	3	0	0,40	1,00	0,57
2	C6	0,060	12960	10	1000	2000	3	2000	2	1	0	0,67	1,00	0,80
	PV18	0,135	12960	10	1000	2000	3	2000	2	2	0	0,50	1,00	0,67
	AD23	0,135	12960	10	1000	2000	3	2000	2	3	0	0,40	1,00	0,57
	LM18	0,100	12960	5	1000	2000	3	2000	2	2	0	0,50	1,00	0,67

Nota: El mejor rendimiento se encuentra en negrita.

6.9.6 Detección de fallas en WT3 mediante Anomaly Transformer

6.9.6.1 Anomaly Transformer utilizando el método RFE 12 en WT3

En la **Figura 28** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías, utilizando las variables del método RFE12. Se puede observar que el modelo logró detectar la primera falla, aproximadamente 7 meses antes de su ocurrencia que fue el 25 de mayo del 2019. Del mismo modo, el modelo advirtió con anticipación sobre

la segunda falla ocurrida el 23 de diciembre de 2019 y tercera falla el 31 de marzo del 2021 hasta con aproximadamente 4 y 6 meses de anticipación, respectivamente.

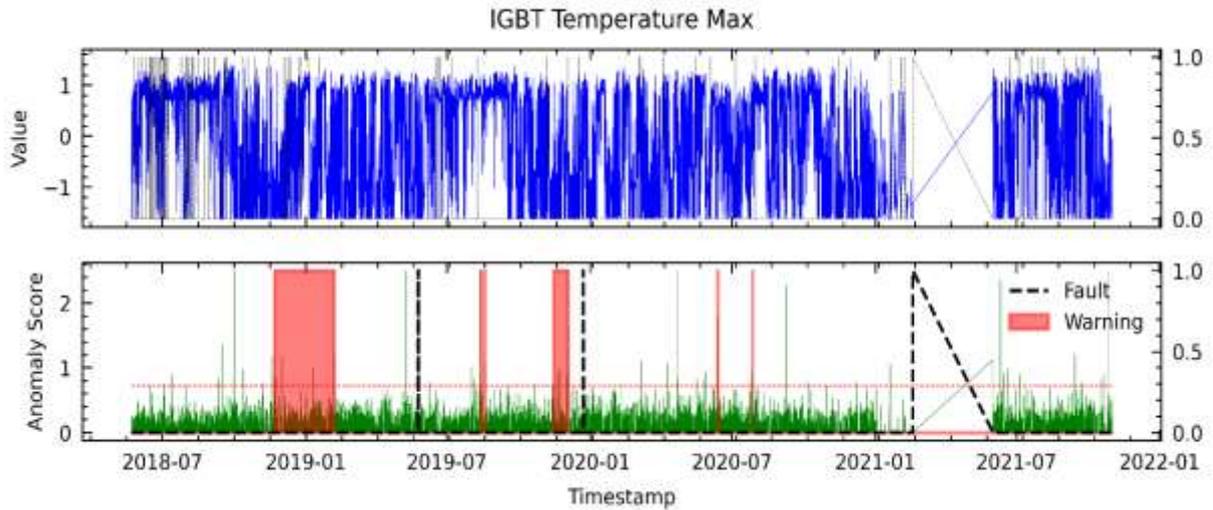


Figura 28. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

6.9.6.2 Resultados de Anomaly Transformer para WT3

Los resultados para la WT3 son presentados en la **Tabla 15**.

Tabla 17. Resultados de rendimiento del modelo AT para WT3.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,040	12960	5	1000	2000	3	2000	3	2	0	0,60	1,00	0,75
	RFE12	0,130	15000	8	5000	2000	3	2000	3	0	0	1,00	1,00	1,00
	RFE24	0,120	15000	8	5000	2000	3	2000	3	2	0	0,60	1,00	0,75
3	C6	0,080	12960	8	5000	2000	3	5000	3	1	0	0,75	1,00	0,86
	PV18	0,150	15000	12	5000	2000	3	2000	3	0	0	1,00	1,00	1,00
	AD23	0,180	15000	8	5000	2000	3	2000	3	1	0	0,75	1,00	0,86
	LM18	0,120	12960	6	5000	2000	3	2000	3	3	0	0,50	1,00	0,67

Nota: El mejor rendimiento se encuentra en negrita.

6.9.7 Detección de fallas en WT4 mediante Anomaly Transformer

6.9.7.1 Anomaly Transformer utilizando el método LM18 en WT4

La **Figura 29** muestra el análisis de la falla del módulo IGBT a través de la puntuación de anomalías utilizando las variables del método LM18. Las fallas ocurridas el 8 de diciembre del 2017 y 04 de marzo de 2020, se puede ver que el modelo logró advertirlas con aproximadamente 8 y 7 meses antes de su ocurrencia respectivamente. Sin embargo, el modelo advirtió de una falla a finales de 2022 que debido a la falta de información después de esa fecha no se puede determinar con certeza si realmente hubo una falla, por lo tanto, se clasifica como un FP.

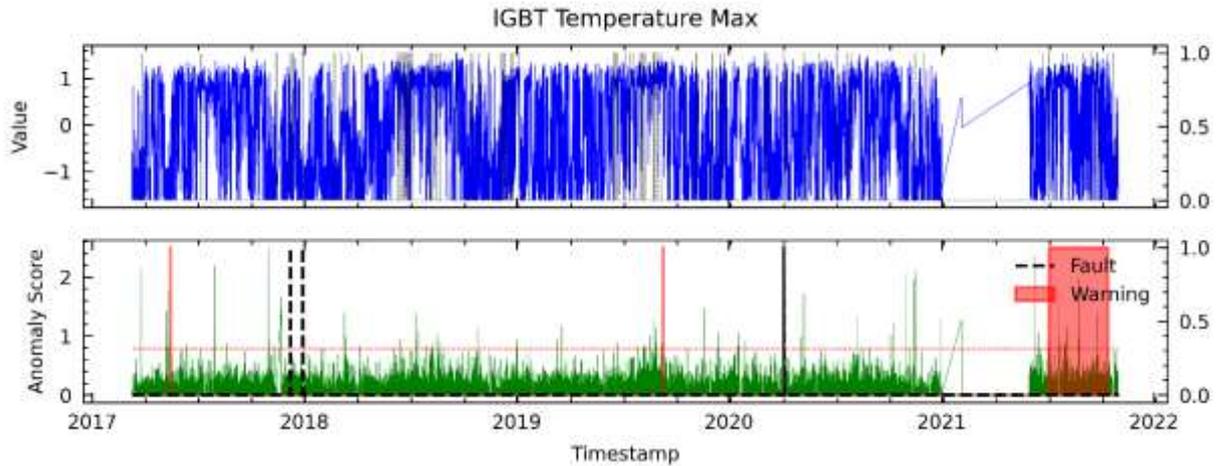


Figura 29. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante modelo AT para WT4.

6.9.7.2 Resultados de Anomaly Transformer para WT4

Los resultados para la WT4 son presentados en la **Tabla 18**.

Tabla 18. Resultados de rendimiento del modelo AT para WT4.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,050	12960	10	5000	2000	4	2000	2	1	0	0,67	1,00	0,80
	RFE12	0,065	12960	9	5000	2000	5	5000	1	0	1	1,00	0,50	0,67
	RFE24	0,060	12960	5	5000	2000	5	2000	2	2	0	0,50	1,00	0,67
4	C6	0,050	15000	7	5000	2000	5	2000	2	2	0	0,50	1,00	0,67
	PV18	0,120	12960	10	5000	2000	4	2000	1	2	1	0,33	0,50	0,40
	AD23	0,075	12960	8	5000	2000	4	2000	1	1	1	0,50	0,50	0,50
	LM18	0,100	12960	10	5000	2000	5	2000	2	1	0	0,67	1,00	0,80

Nota: El mejor rendimiento se encuentra en negrita.

6.9.8 Detección de fallas en WT5 mediante Anomaly Transformer

6.9.8.1 Anomaly Transformer utilizando el método RFE 12 en WT5

En la **Figura 30** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías, utilizando las variables del método RFE12. Se puede observar que, para la falla del 25 de noviembre del 2019, el modelo logró detectarla aproximadamente 3 meses antes de que ocurriera.

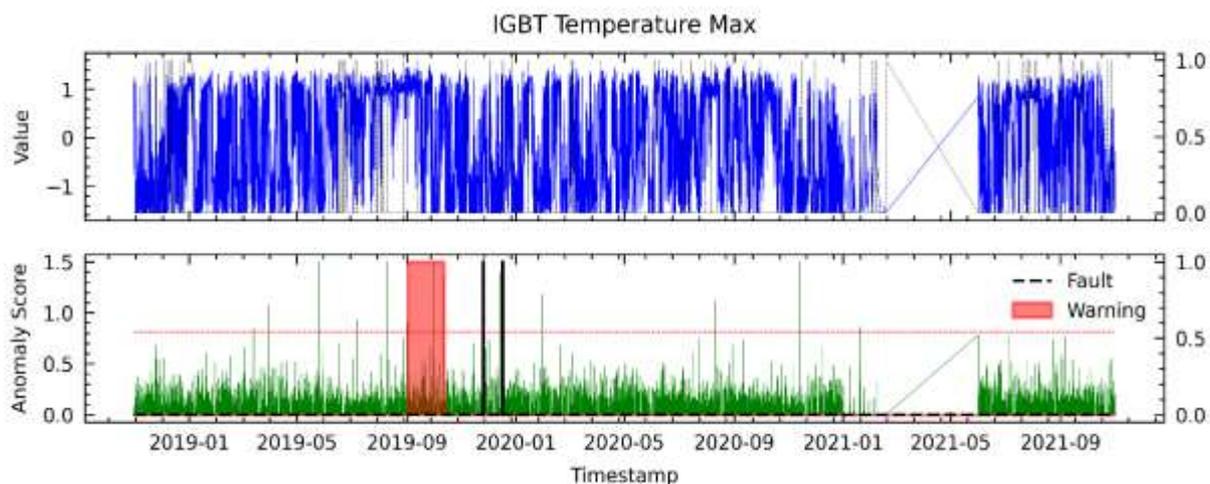


Figura 30. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.

6.9.8.2 Resultados de Anomaly Transformer para WT5

Los resultados para la WT5 presentados en la **Tabla 19**.

Tabla 19. Resultados de rendimiento del modelo AT para WT5.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,050	12960	12	1000	2000	3	2000	1	3	0	0,25	1,00	0,40
	RFE12	0,100	12960	4	1000	2000	3	2000	1	0	0	1,00	1,00	1,00
	RFE24	0,100	12960	4	1000	2000	3	2000	0	1	0	0,00	0,00	0,00
5	C6	0,080	10000	1000	1000	2000	3	2000	1	3	0	0,25	1,00	0,40
	PV18	0,210	12960	7	1000	2000	3	2000	1	2	0	0,33	1,00	0,50
	AD23	0,135	12960	5	1000	2000	3	2000	1	0	0	1,00	1,00	1,00
	LM18	0,120	12960	5	1000	2000	3	2000	0	1	1	0,00	0,00	0,00

Nota: El mejor rendimiento se encuentra en negrita.

6.9.9 Detección de fallas en WT6 mediante Anomaly Transformer

6.9.9.1 Anomaly Transformer utilizando el método AD23 en WT6

En la **Figura 31** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías, utilizando las variables del método AD23. Se puede observar que el modelo logró detectar la falla 3 meses antes de su ocurrencia aproximadamente, que tuvo lugar el 10 de diciembre del 2018. Este hecho fue detectado debido a que antes de la ocurrencia de la falla se comenzó a observar anomalías de manera densa en la serie temporal.

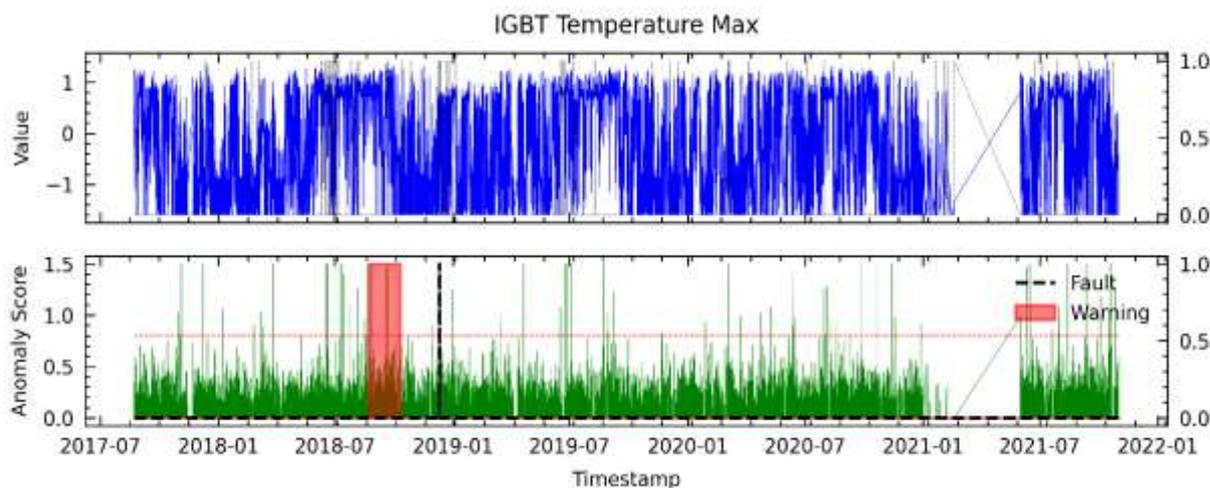


Figura 31. Resultados para AD23 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

6.9.9.2 Resultados de Anomaly Transformer para WT6

Los resultados para la WT6 son presentados en la **Tabla 20**.

Tabla 20. Resultados de rendimiento del modelo AT para WT6.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,020	12960	4	1000	2000	3	2000	1	2	0	0,33	1,00	0,50
	RFE12	0,080	12960	10	1000	2000	20	2000	0	1	1	0,00	0,00	0,00
	RFE24	0,050	12960	6	1000	2000	3	2000	0	2	1	0,00	0,00	0,00
6	C6	0,050	12960	10	5000	2000	20	2000	1	2	0	0,33	1,00	0,50
	PV18	0,135	12960	3	1000	2000	3	2000	1	1	0	0,50	1,00	0,67
	AD23	0,135	12960	10	1000	2000	6	2000	1	0	0	1,00	1,00	1,00
	LM18	0,140	12960	3	1000	2000	6	2000	1	1	0	0,50	1,00	0,67

Nota: El mejor rendimiento se encuentra en negrita.

6.9.10 Detección de fallas en WT7 mediante Anomaly Transformer

6.9.10.1 Anomaly Transformer utilizando el método RFE 12 en WT7

La **Figura 32** muestra el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método RFE12. Se puede observar que el modelo logró detectar la primera falla aproximadamente 7 meses antes de su ocurrencia, que tuvo lugar el 17 de octubre del 2018. Del mismo modo, el modelo advirtió con anticipación sobre la segunda falla y tercera falla ocurridas el 05 de marzo del 2020 y 07 de febrero del 2021, con 6 meses de anticipación aproximadamente. También se registró un FP, porque no se encuentra alguna falla dentro de la ventana de tiempo t_f .

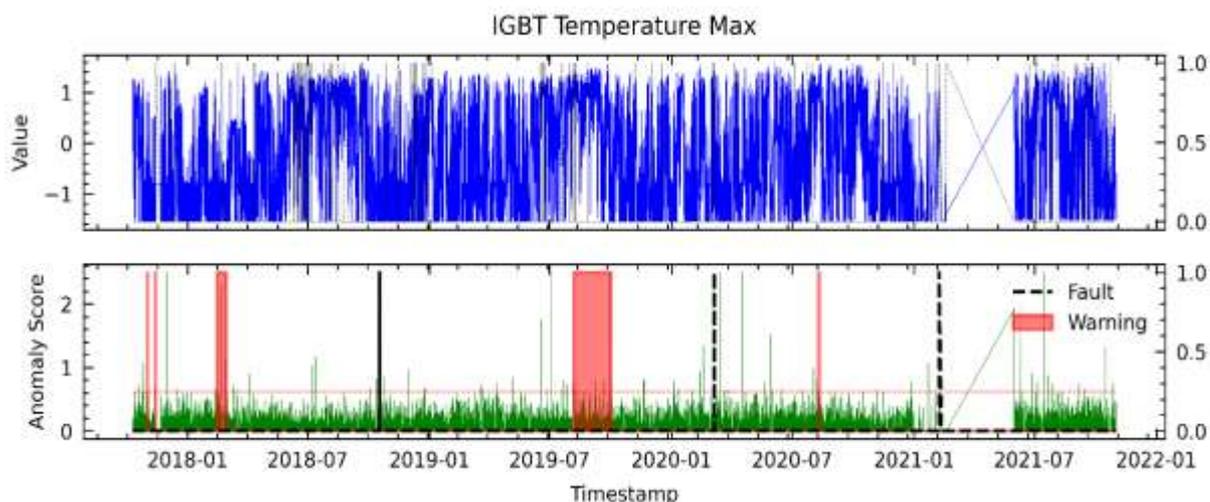


Figura 32. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

6.9.10.2 Resultados de Anomaly Transformer para WT7

Los resultados para la WT7 son presentados en la **Tabla 21**.

Tabla 21. Resultados de rendimiento del modelo AT para WT7.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,040	12960	5	1000	2000	3	2000	3	3	0	0,50	1,00	0,67
	RFE12	0,130	15000	8	5000	2000	3	2000	3	1	0	0,75	1,00	0,86
	RFE24	0,120	15000	8	5000	2000	3	2000	3	3	0	0,50	1,00	0,67
7	C6	0,100	12960	8	5000	2000	3	5000	3	2	0	0,60	1,00	0,75
	PV18	0,135	12960	5	5000	2000	3	2000	3	1	0	0,75	1,00	0,86
	AD23	0,145	12960	5	5000	2000	3	2000	3	3	0	0,50	1,00	0,67
	LM18	0,100	12960	6	5000	2000	3	2000	3	1	0	0,75	1,00	0,86

Nota: El mejor rendimiento se encuentra en negrita.

6.9.11 Detección de fallas en WT8 mediante Anomaly Transformer

6.9.11.1 Anomaly Transformer utilizando el método C6 en WT8

En la **Figura 33** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método C6. Se puede observar que el modelo logró detectar la falla del 19 de febrero del 2018, aproximadamente 2 meses antes de su ocurrencia, ya que advirtió con 2 alarmas que tuvieron un desfase aproximado de 1 mes entre ellas. Del mismo modo, el modelo advirtió una falla en el motor de pitch, la misma que no corresponde al componente estudiado por lo que se considera FP.

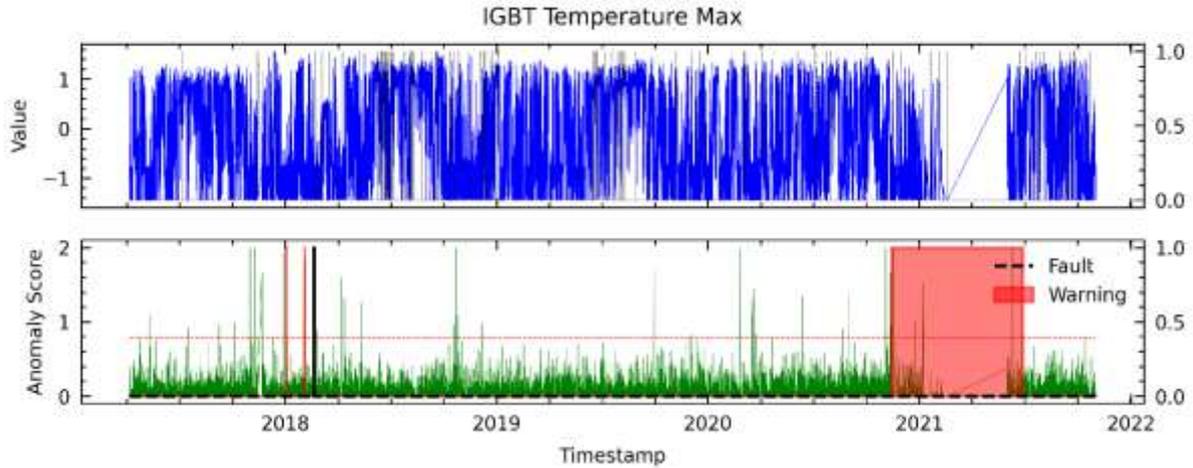


Figura 33. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.

6.9.11.2 Resultados de Anomaly Transformer para WT8

Los resultados para la WT8 son presentados en la **Tabla 22**.

Tabla 22. Resultados de rendimiento del modelo AT para WT8.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,020	12960	8	1000	2000	6	2000	1	0	0	1,00	1,00	1,00
	RFE12	0,108	12960	4	1000	2000	6	2000	0	2	1	0,00	0,00	0,00
	RFE24	0,025	15000	4	1000	2000	6	2000	0	2	1	0,00	0,00	0,00
8	C6	0,021	12960	6	1000	2000	6	2000	1	1	0	0,50	1,00	0,67
	PV18	0,135	12960	8	1000	2000	6	2000	0	3	1	0,00	0,00	0,00
	AD23	0,080	12960	5	1000	2000	6	2000	0	1	2	0,00	0,00	0,00
	LM18	0,023	15000	4	1000	2000	7	2000	0	2	2	0,00	0,00	0,00

Nota: El mejor rendimiento se encuentra en negrita.

6.9.12 Detección de fallas en WT9 mediante Anomaly Transformer

6.9.12.1 Anomaly Transformer utilizando el método AD23 en WT9

En la **Figura 34** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método AD23. Se puede observar que el modelo logró detectar la falla 9 meses antes de su ocurrencia aproximadamente, que fue el 28 de agosto del 2019. También se observa que al momento de advertir la falla, existen múltiples anomalías superan el umbral r , por lo que se advierte al usuario de posible falla en el módulo IGBT.

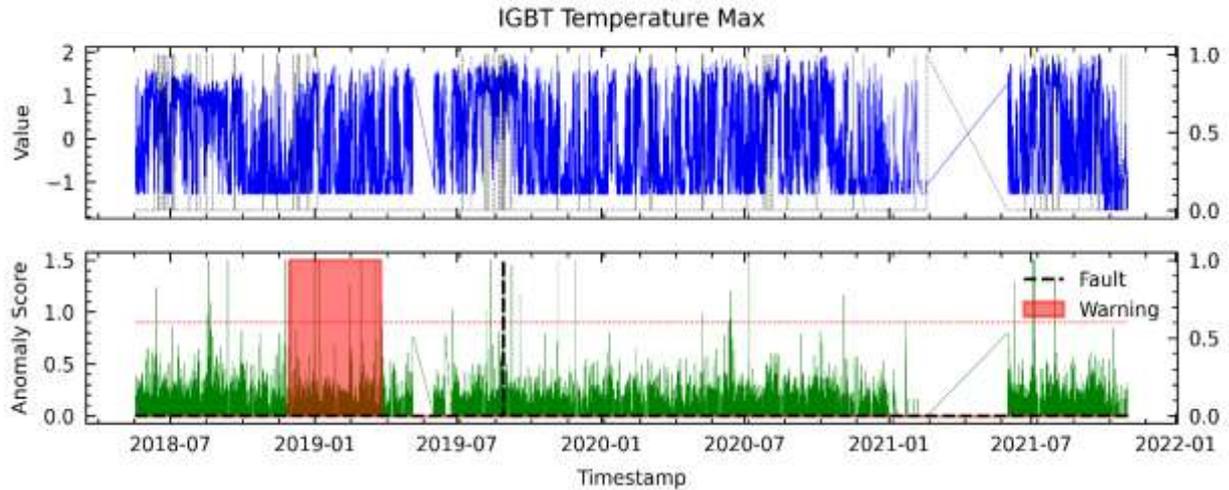


Figura 34. Resultados para AD23 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.

6.9.12.2 Resultados de Anomaly Transformer para WT9

Los resultados para la WT9 son presentados en la **Tabla 23**.

Tabla 23. Resultados de rendimiento del modelo AT para WT9.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,040	12960	5	1000	2000	3	2000	1	4	0	0,20	1,00	0,33
	RFE12	0,175	10000	5	1000	2000	6	2000	1	1	0	0,50	1,00	0,67
	RFE24	0,075	12960	8	1000	2000	6	2000	1	2	0	0,33	1,00	0,50
9	C6	0,080	12960	12	1000	2000	6	2000	1	1	0	0,50	1,00	0,67
	PV18	0,160	12960	3	1000	2000	3	2000	1	2	0	0,33	1,00	0,50
	AD23	0,135	12960	5	1000	2000	3	2000	1	0	0	1,00	1,00	1,00
	LM18	0,120	15000	11	1000	2000	6	2000	1	1	0	0,50	1,00	0,67

Nota: El mejor rendimiento se encuentra en negrita.

6.9.13 Detección de fallas en WT10 mediante Anomaly Transformer

6.9.13.1 Anomaly Transformer utilizando el método C6 en WT10

En la **Figura 35** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método C6. Se puede observar que el modelo logró advertir las fallas del 17 de junio del 2019 y del 17 de septiembre del 2021 con aproximadamente 6 y 4 meses respectivamente. En este caso, también se puede observar que la alarma se mantienen hasta 1 mes antes de la falla. Finalmente, se considera que las dos últimas fallas registradas se encuentran a menos de 1 mes del reemplazo del componente, por lo que el modelo no puede capturar anomalías en casos fortuitos.

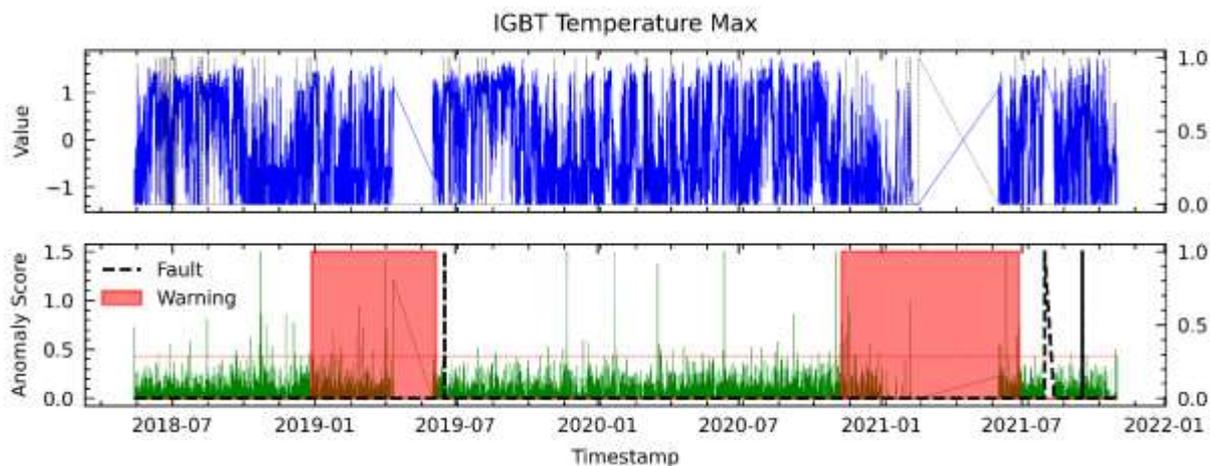


Figura 35. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante modelo AT para WT10.

6.9.13.2 Resultados de Anomaly Transformer para WT10

Los resultados para la WT10 son presentados en la **Tabla 24**.

Tabla 24. Resultados de rendimiento del modelo AT para WT10.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,050	12960	13	1000	2000	6	2000	2	0	0	1,00	1,00	1,00
	RFE12	0,100	12960	4	1000	2000	6	2000	1	1	1	0,50	0,50	0,50
	RFE24	0,100	12960	12	1000	2000	6	2000	2	0	0	1,00	1,00	1,00
10	C6	0,080	12960	12	1000	2000	6	2000	2	0	0	1,00	1,00	1,00
	PV18	0,150	12960	3	1000	2000	6	2000	2	1	0	0,67	1,00	0,80
	AD23	0,135	12960	6	1000	2000	6	2000	2	1	0	0,67	1,00	0,80
	LM18	0,100	12960	8	1000	2000	6	2000	2	1	0	0,67	1,00	0,80

Nota: El mejor rendimiento se encuentra en negrita.

6.9.14 Detección de fallas en WT11 mediante Anomaly Transformer

6.9.14.1 Anomaly Transformer utilizando el método C6 en WT11

La **Figura 36** muestra el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método C6. Se puede observar que el modelo logró detectar la falla aproximadamente 7 meses antes de su ocurrencia, que tuvo lugar el 03 de agosto de 2021. No se puede contrastar más resultados puesto que no hubo otra falla registrada en este periodo.

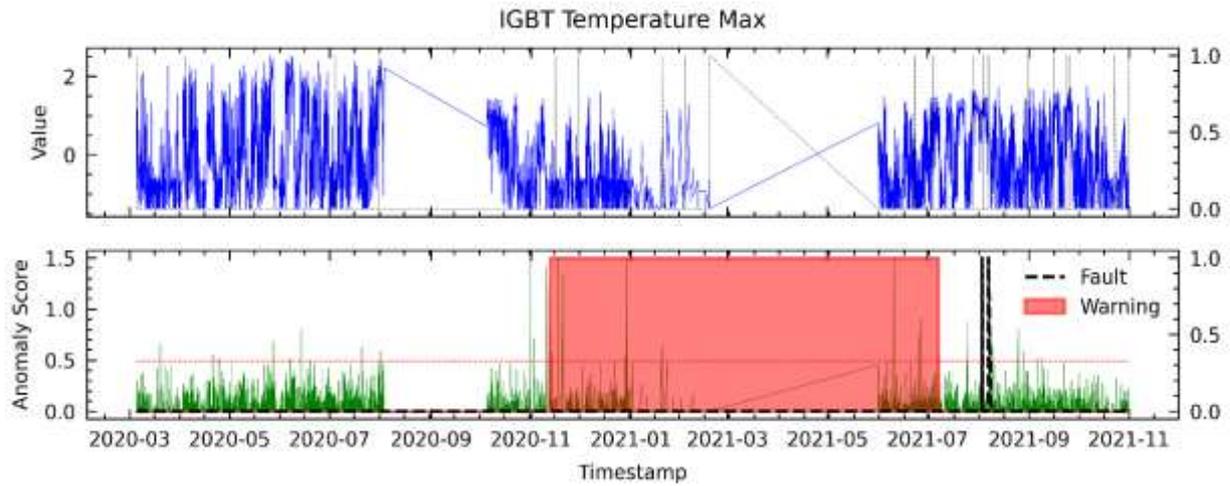


Figura 36. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

6.9.14.2 Resultados de Anomaly Transformer para WT11

Los resultados para la WT11 son presentados en la **Tabla 25**.

Tabla 25. Resultados de rendimiento del modelo AT para WT11.

WTG	FS	r	N	thresh	a	N2	thresh2	a2	TP	FP	FN	RC	PR	F1
	RFE1	0,025	5000	3	1000	2000	6	2000	1	2	0	0,33	1,00	0,50
	RFE12	0,135	6000	4	1000	2000	6	2000	1	1	0	0,50	1,00	0,67
	RFE24	0,050	6000	4	1000	2000	6	2000	1	1	0	0,50	1,00	0,67
11	C6	0,060	5000	4	1000	2000	6	2000	1	0	0	1,00	1,00	1,00
	PV18	0,150	5000	4	1000	2000	6	2000	1	0	0	1,00	1,00	1,00
	AD23	0,135	6000	3	1000	2000	6	2000	1	1	0	0,50	1,00	0,67
	LM18	0,150	5000	4	1000	2000	6	2000	1	1	0	0,50	1,00	0,67

Nota: El mejor rendimiento se encuentra en negrita.

6.10 Aplicación del segundo modelo de detección de anomalías TranAD

En este apartado se evaluaron los resultados de cuatro métodos de FS: RF12, C6, LM18 y PV18. Debido a las limitaciones computacionales, solo se pudo procesar hasta 18 variables, por lo que los otros métodos de selección de variables se recomiendan que sean abordados en una investigación futura.

6.10.1 Ventana deslizante para detección de fallas (segunda etapa)

La evaluación del modelo TranAD se realizó en dos etapas, teniendo en cuenta que estas dos etapas pueden ser reguladas según el operario para ajustarse a las fallas del módulo IGBT. En la primera etapa, se realizó el mismo proceso usado en la evaluación de fallas de la **sección 6.9.2** para el modelo Anomaly Transformer, en este caso se lo realizó para cada variable. En la segunda etapa, se tomaron las etiquetas de anomalías de cada variable y se sumaron para obtener un solo vector. Luego, se aplicó un nuevo umbral *threshEt2* a este vector. Si las anomalías se repiten superando el umbral de densidad de anomalías *tw* en un periodo de tiempo *n3*, serán etiquetadas y enviadas como advertencias de falla del módulo IGBT. Cabe recalcar que los valores de los parámetros son diferentes para cada método de selección de características y cada WT evaluado. Estos cambios pueden deberse a diferentes características seleccionadas, así como de la naturaleza inherentemente variable de las condiciones del viento.

6.10.2 Detección de fallas en WT1 mediante TranAD

6.10.2.1 TranAD utilizando el método RF12 en WT1

En este método se analiza de manera independiente variable por variable, donde se aplica la evaluación de primera etapa, se observa en la **Figura 37**, cada variable, su puntaje de anomalías y la advertencia de posible falla.

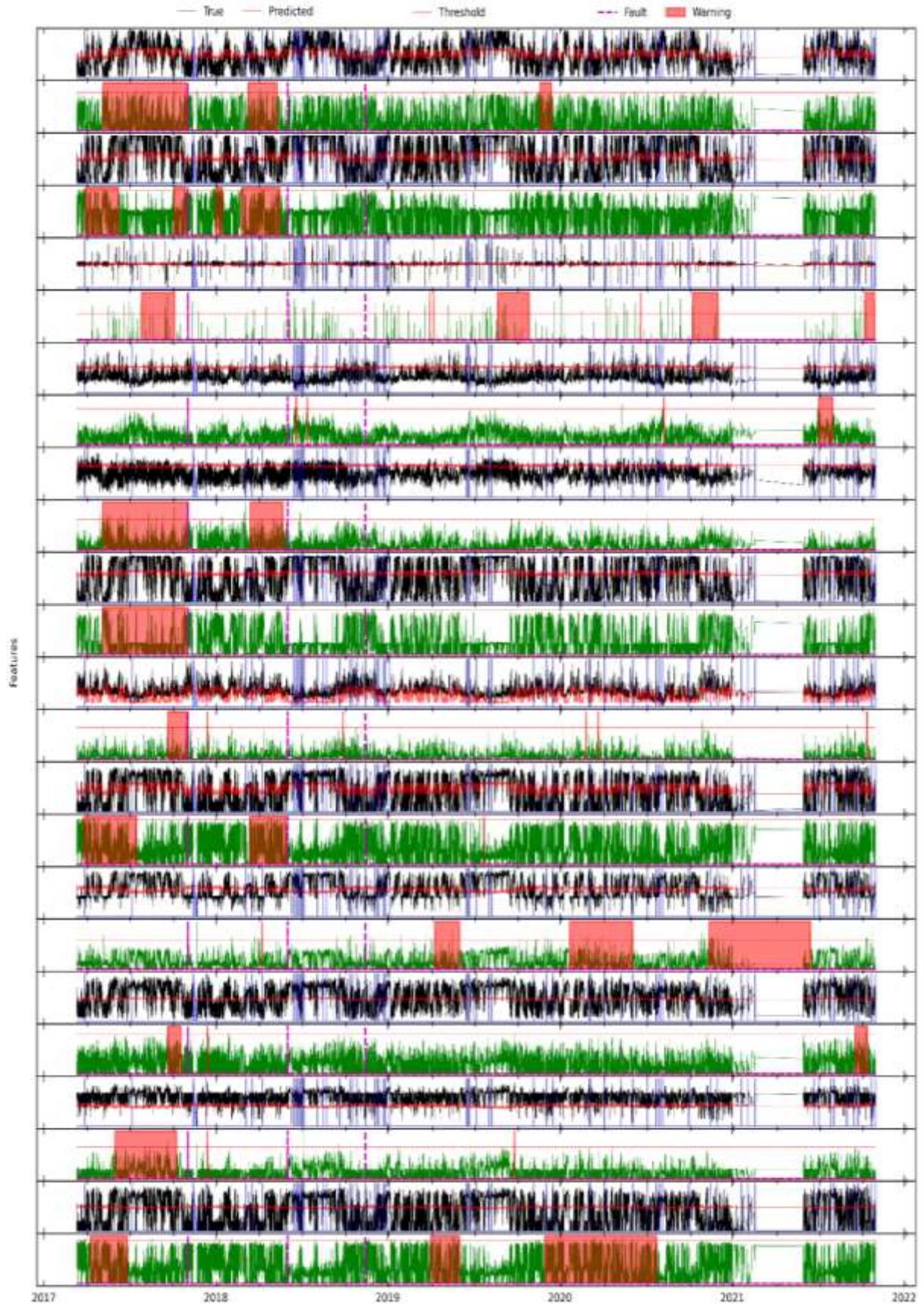


Figura 37. Resultados para RFE12 en el análisis de evaluación de primera etapa mediante el modelo TranAD para WT1.

La evaluación de resultados de segunda etapa se observa en la **Figura 38**. Aquí el modelo predijo la primera falla ocurrida el 02 de noviembre de 2017, aproximadamente 6 meses antes; y la segunda ocurrida el 02 de junio de 2018 con aproximadamente 3 meses antes; mientras que con la tercera que tuvo lugar el 14 de noviembre de 2018 no se pudo predecir.

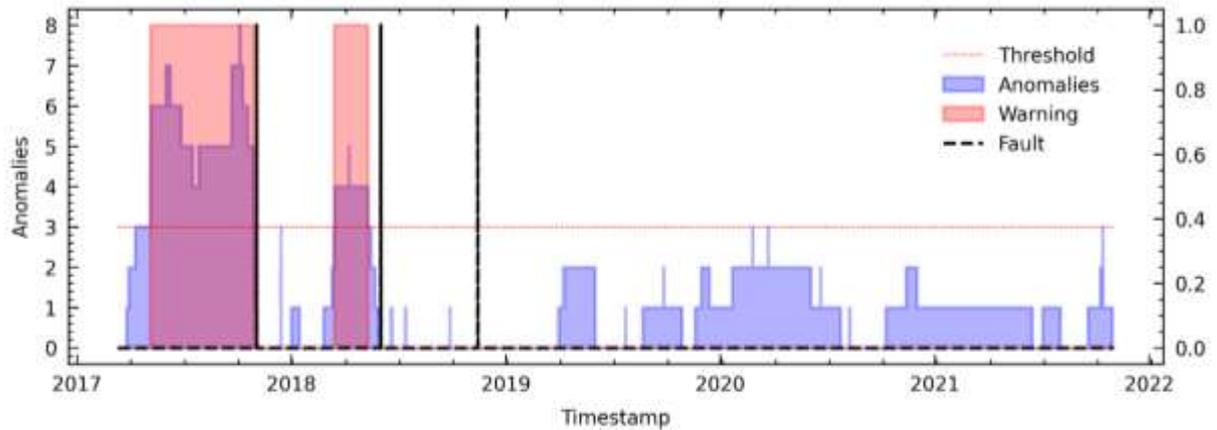


Figura 38. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.

6.10.2.2 TranAD utilizando el método C6 en WT1

De igual manera, como se indicó en el apartado anterior, aquí y en los apartados subsiguientes, se realizó la evaluación de primera etapa para obtener el puntaje de anomalías y la advertencia de posible falla. El análisis de segunda etapa se observa en la **Figura 39**, en la que se destaca que el modelo predijo las tres fallas en el módulo IGBT con un promedio de 2 meses de anticipación aproximadamente. Cada falla fue anticipada con aproximadamente 3, 3 y 1 meses respectivamente. Si se desea observar con mayor detalle, este análisis y los subsiguientes se presentan en el **Anexo 4**.

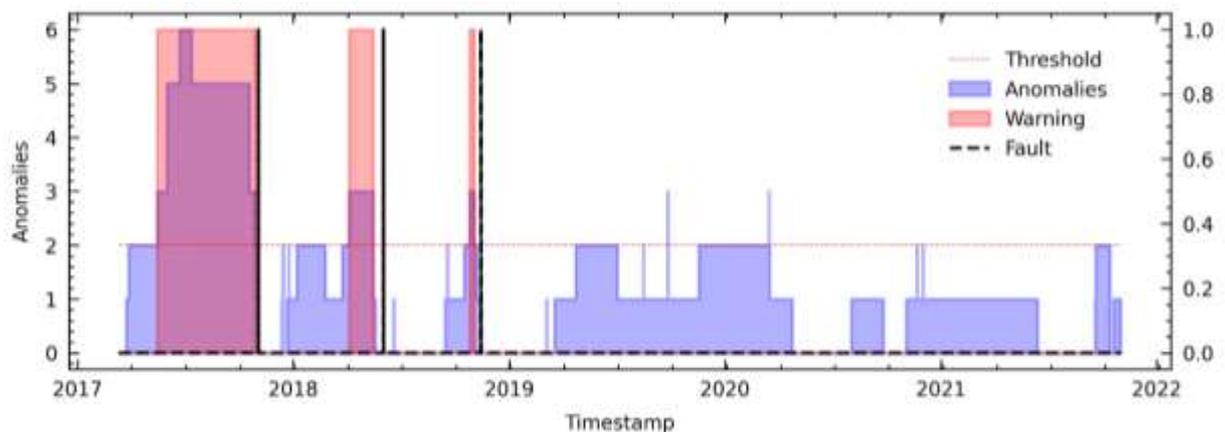


Figura 39. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.

6.10.2.3 TranAD utilizando el método PV18 en WT1

En la **Figura 40** se muestra que el modelo solo logró predecir la primera falla en el módulo IGBT con 3 meses de anticipación aproximadamente, por lo que en este caso se observó dos FNs.

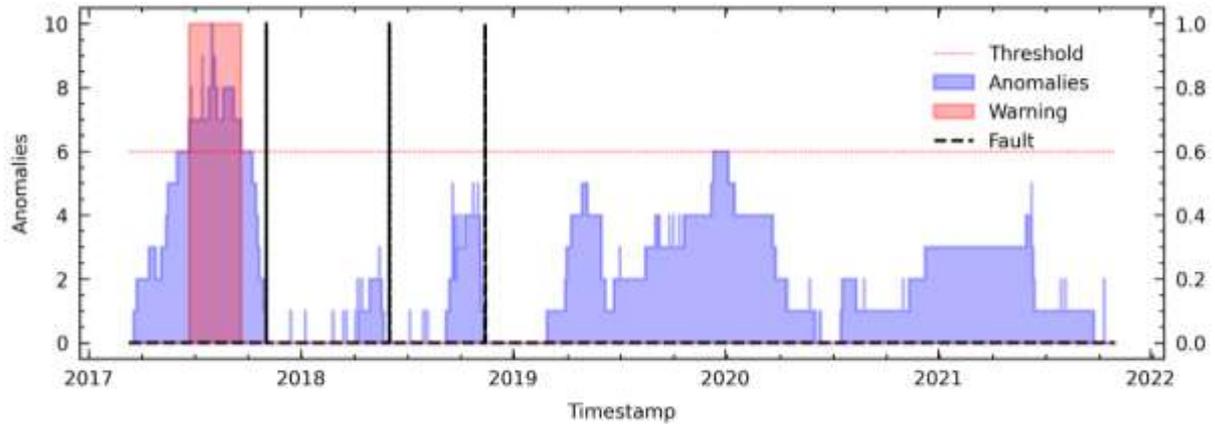


Figura 40. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.

6.10.2.4 TranAD utilizando el método LM18 en WT1

En la **Figura 41** donde se observa que el modelo advirtió la primera falla hasta 6 meses antes de su ocurrencia aproximadamente, así mismo la segunda falla fue predicha aproximadamente 2 meses antes de que ocurra, pero en cambio no logró predecir la tercera falla por lo que se obtuvo un FN. Esto puede deberse a que la última falla produjo un error en la frecuencia del módulo IGBT, por lo que se recomienda su análisis en una investigación futura.

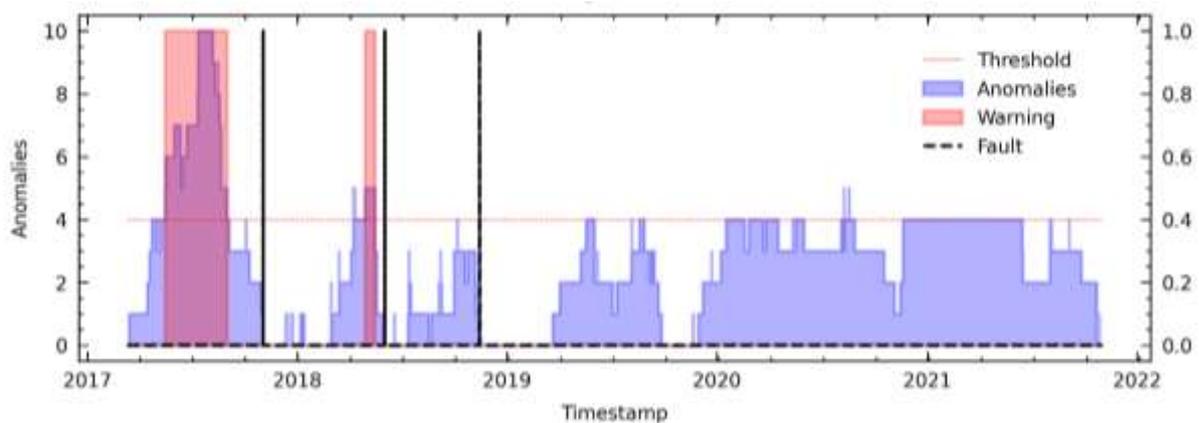


Figura 41. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.

6.10.2.5 Resultados de TranAD para WT1

Los resultados para la WT1 son presentados en la **Tabla 26**, donde se aplicaron los siguientes hiper-parámetros: tamaño de ventana de 12, un paso de 1, 5 épocas de entrenamiento y una tasa de aprendizaje de $1e^{-6}$. Los parámetros *thresh2*, *a2* se los fija en 6 y 2000 respectivamente. Estos hiper- parámetros y parámetros son utilizados para todos los WTs.

Tabla 26. Resultados de rendimiento del modelo TranAD para WT1.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
1	RFE12	0,050	12960	7	1000	2000	4	3	4	2	0	1	1,00	0,667	0,80
	C6	0,100	12960	7	1000	2000	4	2	4	3	0	0	1,00	1,00	1,00
	PV18	0,050	12960	7	1000	2000	4	4	4	1	0	2	1,00	0,33	0,50
	LM18	0,050	12960	7	1000	2000	4	6	4	2	0	1	1,00	0,67	0,80

Nota: El mejor rendimiento se encuentra en negrita.

Para evaluación de primera etapa *r* es el umbral percentil desde donde se toman anomalías utilizado en dicho método, *N* y *N2* son las ventanas de datos utilizadas para contar anomalías, *thresh* y *thresh2* es el umbral que debe superarse para mostrar una advertencia de fallo en el módulo IGBT, *a* y *a2* son los rangos de datos utilizados para determinar si se encuentra otra advertencia correspondiente a la misma falla. Para la evaluación de segunda etapa, *N3* es el tamaño de la ventana de densidad de anomalías, *tw* es el umbral de densidad de anomalías, *threshEt2* es el umbral de variables desde donde se consideran anomalías.

6.10.3 Detección de fallas en WT2 mediante TranAD

A continuación, se presentan los resultados con la mejor puntuación de F1 Score con el objetivo de presentarle al lector los mejores resultados de los métodos desarrollados en la investigación. Si se requiere más detalles se sugiere revisar el **Anexo 4**.

6.10.3.1 TranAD utilizando el método PV18 en WT2

En la **Figura 42** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método PV18. Se puede observar que el modelo logró detectar las dos fallas con promedio de 7 meses de anticipación aproximadamente. Un fallo en el módulo IGBT puede causar anomalías simultáneas en varias variables, lo que puede ser resultado de la degradación del módulo debido a la exposición a altas temperaturas o al uso prolongado. Estos fallos pueden causar daños graves en el sistema, reducir la eficiencia y la durabilidad de los aerogeneradores. Se debe seguir investigando acerca de las posibles fallas de este componente más a fondo, para determinar con más exactitud que variables están comprometidas. En las fallas ocurridas el 10 de junio de 2018 y 06 de marzo de 2021 el modelo advirtió esta falla aproximadamente 8 y 6 meses respectivamente.

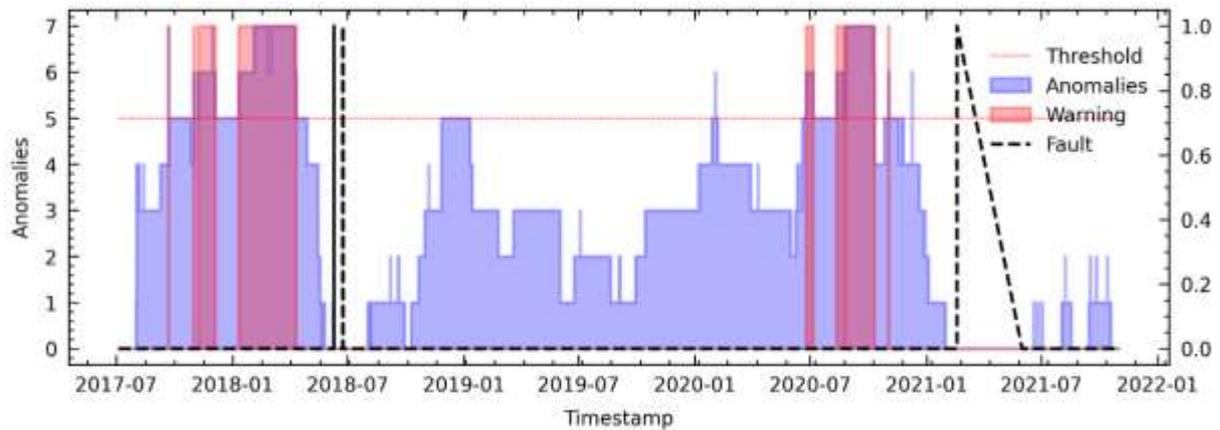


Figura 42. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.

6.10.3.2 Resultados de TranAD para WT2

Los resultados para la WT2 son presentados en la **Tabla 27**.

Tabla 27. Resultados de rendimiento del modelo TranAD para WT3.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	a2	TP	FP	FN	RC	PR	F1
2	RFE12	0,050	12960	7	1000	2000	4	4	4	2	1	0	0,67	1,00	0,80
	C6	0,050	12960	7	1000	2000	4	4	4	1	1	1	0,50	0,50	0,50
	PV18	0,050	12960	7	1000	2000	4	5	4	2	0	0	1,00	1,00	1,00
	LM18	0,050	12960	7	1000	2000	4	4	4	2	3	0	0,40	1,00	0,57

Nota: El mejor rendimiento se encuentra en negrita.

6.10.4 Detección de fallas en WT3 mediante TranAD

6.10.4.1 TranAD utilizando el método PV18 en WT3

En la **Figura 43** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías, utilizando las variables del método PV18. Se puede observar que el modelo logró detectar las fallas con un promedio aproximado de 7 meses antes de su ocurrencia sin dar ningún FP. En la primera falla ocurrida el 25 de mayo de 2019, el modelo advirtió y se mantuvo en alarma 10 meses antes aproximadamente, para la segunda falla el 23 de diciembre de 2019 el modelo la advirtió aproximadamente 3 meses antes. Finalmente, la tercera falla el modelo activo la alarma un aproximado de 7 meses antes de que esta ocurriese el 31 de marzo del 2021.

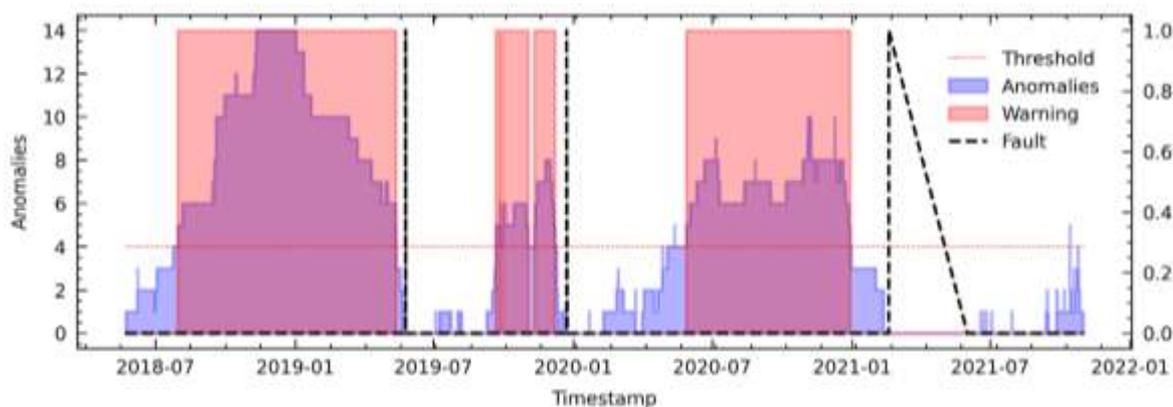


Figura 43. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.

6.10.4.2 Resultados de TranAD para WT3

Los resultados para la WT3 son presentados en la **Tabla 28**.

Tabla 28. Resultados de rendimiento del modelo TranAD para WT3.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
3	RFE12	0,100	12960	7	1000	2000	4	3	4	3	1	0	0,75	1,00	0,86
	C6	0,100	12960	7	1000	2000	4	2	4	3	1	0	0,75	1,00	0,86
	PV18	0,100	12960	7	1000	2000	4	4	4	3	0	0	1,00	1,00	1,00
	LM18	0,100	12960	7	1000	2000	4	3	4	3	1	0	0,75	1,00	0,86

Nota: El mejor rendimiento se encuentra en negrita.

6.10.5 Detección de fallas en WT4 mediante TranAD

6.10.5.1 TranAD utilizando el método PV18 en WT4

En la **Figura 44** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías, utilizando las variables del método PV18. Se puede observar que el modelo logró detectar las fallas con promedio de 5 meses antes de su ocurrencia, esto debido a que hubo una acumulación de anomalías antes de las fallas. En la primera, que ocurrió el 08 de diciembre de 2017, el puntaje de anomalías superó el umbral y advirtió en un aproximado de 6 meses antes de que ocurra la falla. En la segunda falla, el modelo advirtió 3 meses antes aproximadamente de que ocurra el 04 de marzo de 2020.

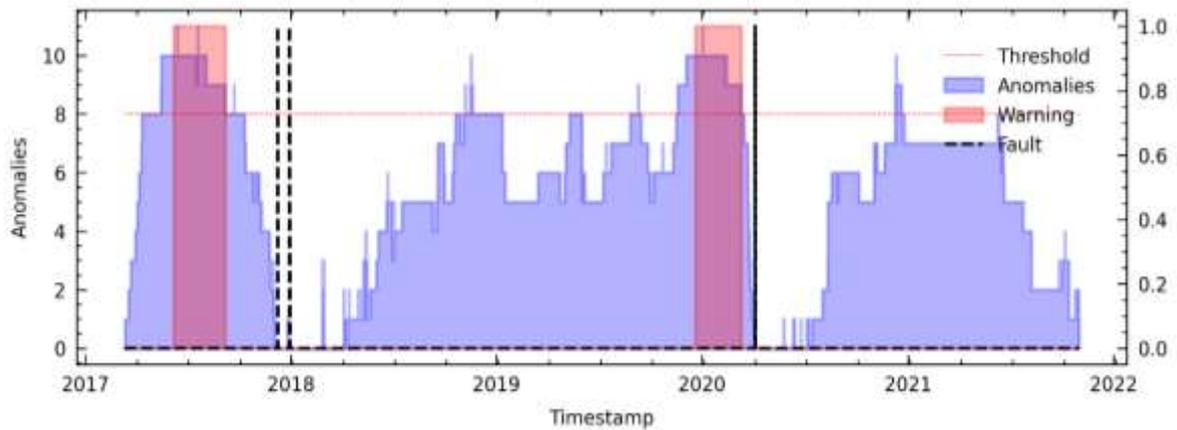


Figura 44. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.

6.10.5.2 Resultados de TranAD para WT4

Los resultados para la WT4 son presentados en la **Tabla 29**.

Tabla 29. Resultados de rendimiento del modelo TranAD para WT4.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
	RFE12	0,050	12960	7	1000	2000	4	3	4	1	0	1	1,00	0,50	0,67
4	C6	0,050	12960	7	1000	2000	4	4	4	1	0	1	1,00	0,50	0,67
	PV18	0,100	12960	7	1000	2000	4	4	4	2	0	0	1,00	1,00	1,00
	LM18	0,050	12960	7	1000	2000	4	4	4	2	3	0	0,40	1,00	0,57

Nota: El mejor rendimiento se encuentra en negrita.

6.10.6 Detección de fallas en WT5 mediante TranAD

6.10.6.1 TranAD utilizando el método RFE12 en WT5

En la **Figura 45** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método RFE12. Se puede observar que el modelo logró predecir la falla del 25 de noviembre de 2019 por una acumulación de anomalías con aproximadamente 3 meses antes de su ocurrencia.

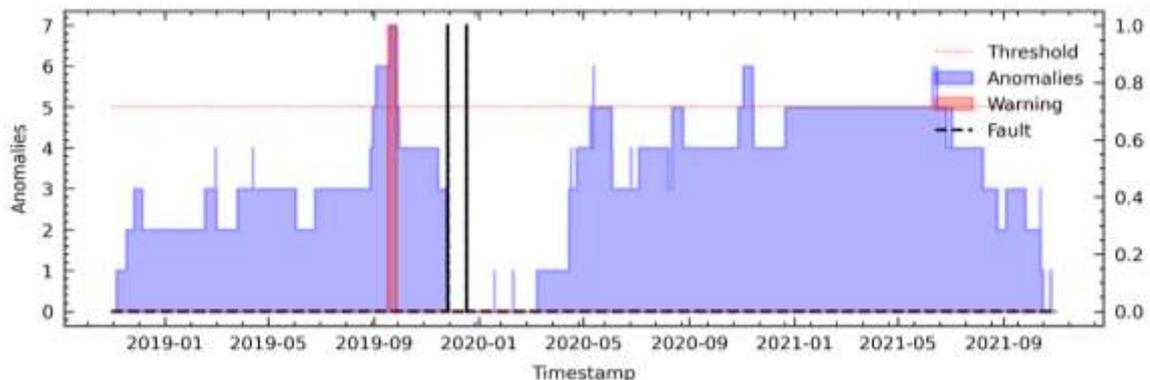


Figura 45. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.

6.10.6.2 Resultados de TranAD para WT5

Los resultados para la WT5 son presentados en la **Tabla 30**.

Tabla 30. Resultados de rendimiento del modelo TranAD para WT5.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
5	RFE12	0,075	12960	7	1000	2000	1500	5	1500	1	0	0	1,00	1,00	1,00
	C6	0,100	12960	7	1000	2000	2000	2	2000	1	1	0	0,50	1,00	0,67
	PV18	0,070	12960	7	1000	2000	3000	8	3000	0	1	1	0,00	0,00	0,00
	LM18	0,050	12960	7	1000	2000	3000	6	3000	0	1	1	0,00	0,00	0,00

Nota: El mejor rendimiento se encuentra en negrita.

6.10.7 Detección de fallas en WT6 mediante TranAD

6.10.7.1 TranAD utilizando el método C6 en WT6

En la **Figura 46** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método PV18. Se puede observar que el modelo predijo la falla aproximadamente 12 meses antes, y esta se mantiene presente durante 3 meses. Además, el modelo también predijo de la falla aproximadamente 2 meses antes de que ocurra.

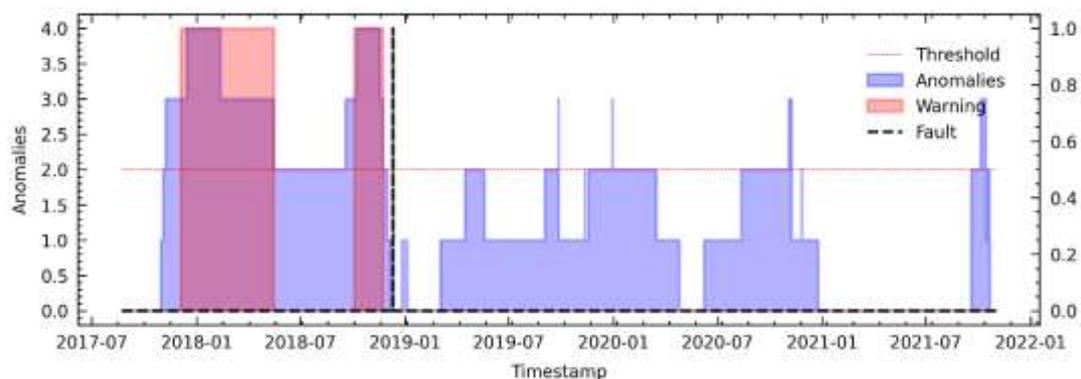


Figura 46. Resultados para C6 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.

6.10.7.2 Resultados de TranAD para WT6

Los resultados para la WT6 son presentados en la **Tabla 31**.

Tabla 31. Resultados de rendimiento del modelo TranAD para WT6.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
6	RFE12	0,075	12960	7	1000	2000	2000	4	2000	1	1	0	0,50	1,00	0,67
	C6	0,100	12960	7	1000	2000	2000	3	2000	1	0	0	1,00	1,00	1,00
	PV18	0,075	12960	7	1000	2000	2000	7	2000	1	0	0	1,00	1,00	1,00
	LM18	0,050	12960	7	1000	2000	2000	4	2000	1	2	0	0,33	1,00	0,50

Nota: El mejor rendimiento se encuentra en negrita.

6.10.8 Detección de fallas en WT7 mediante TranAD

6.10.8.1 TranAD utilizando el método LM18 en WT7

En la **Figura 47** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método LM18. Se puede observar que el modelo logró detectar las fallas con un promedio aproximado de 5 meses antes de su ocurrencia. En el caso de la primera falla ocurrida el 17 de octubre de 2018, se activó una alarma aproximadamente 12 meses antes de su ocurrencia, y luego 3 meses antes. Para la segunda falla que tuvo lugar el 05 de marzo del 2020, se logró predecir aproximadamente con 2 meses de anticipación. Finalmente, para la tercera falla que fue el 07 de febrero del 2021, se produjeron 2 alarmas seguidas con una diferencia aproximada de un mes, que predijeron la falla con 5 meses de anticipación aproximadamente.

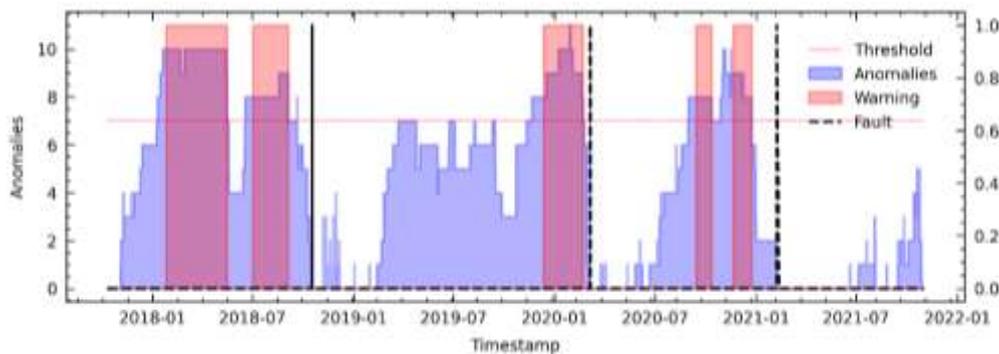


Figura 47. Resultados para LM18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.

6.10.8.2 Resultados de TranAD para WT7

Los resultados para la WT7 son presentados en la **Tabla 32**.

Tabla 32. Resultados de rendimiento del modelo TranAD para WT7.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
7	RFE12	0,100	12960	7	1000	2000	2000	4	2000	3	1	0	0,75	1,00	0,86
	C6	0,100	12960	7	1000	2000	2000	2	2000	3	0	0	1,00	1,00	1,00
	PV18	0,100	12960	7	1000	2000	2000	7	2000	3	1	0	0,75	1,00	0,86
	LM18	0,100	12960	7	1000	2000	2000	7	2000	3	0	0	1,00	1,00	1,00

Nota: El mejor rendimiento se encuentra en negrita.

6.10.9 Detección de fallas en WT8 mediante TranAD

6.10.9.1 TranAD utilizando el método PV18 en WT8

En la **Figura 48** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método RFE12. Se puede observar que el modelo logró predecir la falla aproximadamente con seis meses de anticipación, la cual tuvo lugar el 19 de febrero del 2018.

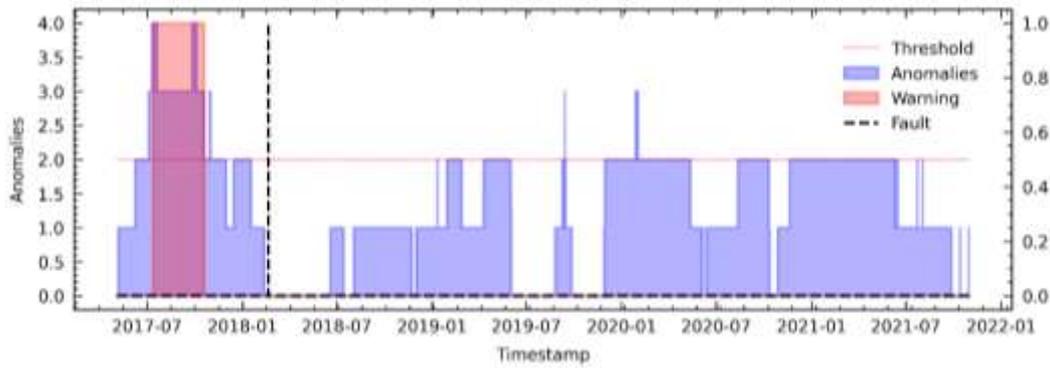


Figura 48. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.

6.10.9.2 Resultados de TranAD para WT8

Los resultados para la WT8 son presentados en la **Tabla 33**.

Tabla 33. Resultados de rendimiento del modelo TranAD para WT8.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
8	RFE12	0,040	12960	7	1000	2000	1000	2	1000	1	0	0	1,00	1,00	1,00
	C6	0,100	12960	7	1000	2000	1000	3	1000	1	1	0	0,50	1,00	0,67
	PV18	0,038	12960	7	1000	2000	1000	9	1000	1	0	0	1,00	1,00	1,00
	LM18	0,075	12960	7	1000	2000	3000	4	3000	1	1	0	0,50	1,00	0,67

Nota: El mejor rendimiento se encuentra en negrita.

6.10.10 Detección de fallas en WT9 mediante TranAD

6.10.10.1 TranAD utilizando el método RFE12 en WT9

En la **Figura 49** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método RFE12. Se puede observar que el modelo logró predecir la falla aproximadamente 10 meses antes de que tenga lugar el 28 de agosto del 2019.

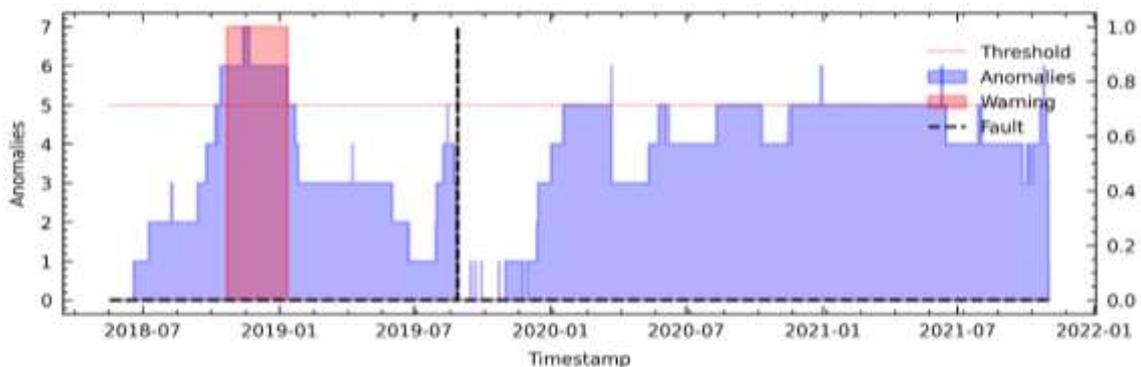


Figura 49. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.

6.10.10.2 Resultados de TranAD para WT9

Los resultados para la WT9 son presentados en la **Tabla 34**.

Tabla 34. Resultados de rendimiento del modelo TranAD para WT9.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
9	RFE12	0,100	12960	7	1000	2000	1000	5	1000	1	0	0	1,00	1,00	1,00
	C6	0,200	12960	7	1000	2000	1000	3	1000	1	2	0	0,33	1,00	0,50
	PV18	0,080	12960	7	1000	2000	1000	7	1000	1	1	0	0,50	1,00	0,67
	LM18	0,080	12960	7	1000	2000	2000	7	2000	1	2	0	0,33	1,00	0,50

Nota: El mejor rendimiento se encuentra en negrita.

6.10.11 Detección de fallas en WT10 mediante TranAD

6.10.11.1 TranAD utilizando el método RFE12 en WT10

En la **Figura 50** se presenta el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método RFE12. Se puede observar que el modelo logró predecir las fallas aproximadamente 7 meses antes de su ocurrencia. En la primera falla, ocurrida el 17 de junio de 2019, el modelo la predijo aproximadamente 8 meses antes, y en la segunda falla logró predecir aproximadamente con 7 meses antes de que suceda el 11 de agosto de 2021.

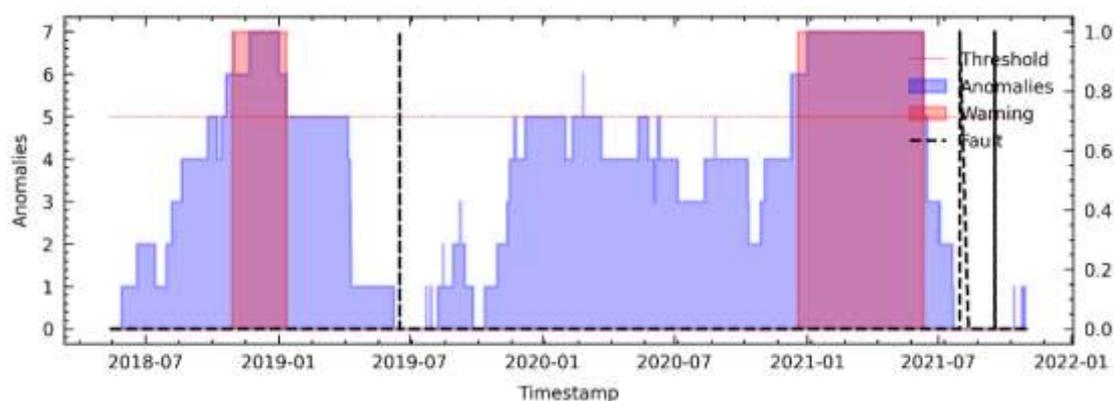


Figura 50. Resultados para RFE12 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.

6.10.11.2 Resultados de TranAD para WT10

Los resultados para la WT10 son presentados en la **Tabla 35**.

Tabla 35. Resultados de rendimiento del modelo TranAD para WT10.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
10	RFE12	0,080	12960	7	1000	2000	1000	5	1000	2	0	0	1,00	1,00	1,00
	C6	0,250	12960	7	1000	2000	1000	3	1000	1	1	1	0,50	0,50	0,50
	PV18	0,100	12960	7	1000	2000	200	7	200	2	1	0	0,67	1,00	0,80
	LM18	0,070	12960	7	1000	2000	360	8	360	2	0	0	1,00	1,00	1,00

Nota: El mejor rendimiento se encuentra en negrita.

6.10.12 Detección de fallas en WT11 mediante TranAD

6.10.12.1 TranAD utilizando el método PV18 en WT11

La **Figura 51** muestra el resultado del análisis de la falla del módulo IGBT a través del puntaje de anomalías utilizando las variables del método PV18. En el caso de la primera falla, el modelo la predijo aproximadamente 6 meses antes de que tuviera lugar el 03 de agosto del 2021. Además, se observa una alarma donde no hubo registro de falla por lo que fue etiquetada como FP.

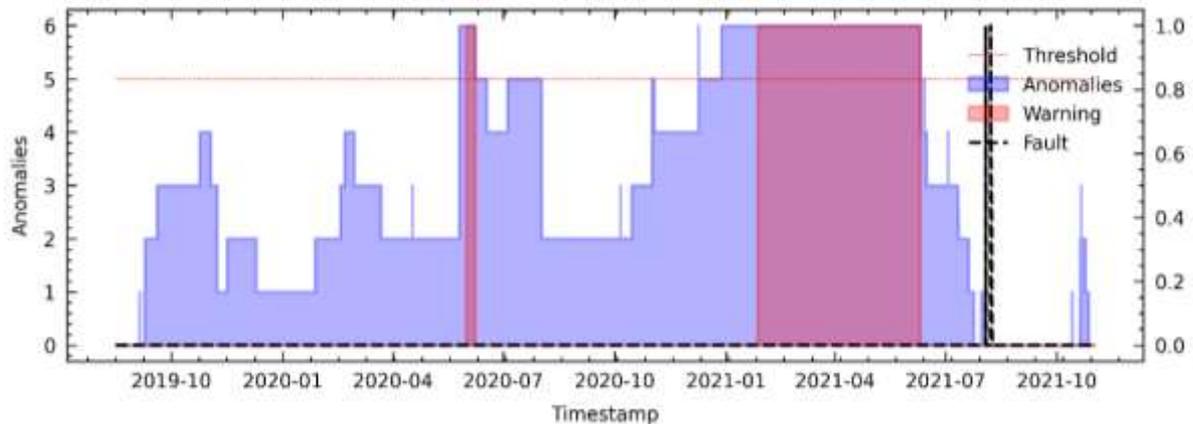


Figura 51. Resultados para PV18 que obtuvo el mejor rendimiento en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.

6.10.12.2 Resultados de TranAD para WT11

Los resultados para la WT11 son presentados en la **Tabla 36**.

Tabla 36. Resultados de rendimiento del modelo TranAD para WT11.

WTG	FS	r	N	thresh	a	N2	N3	threshEt2	tw	TP	FP	FN	RC	PR	F1
11	RFE12	0,100	12960	7	1000	2000	1000	6	1000	0	1	1	0,00	0,00	0,00
	C6	0,100	12960	7	1000	2000	1000	3	1000	0	1	1	0,00	0,00	0,00
	PV18	0,050	12960	7	1000	2000	500	5	500	1	1	0	0,50	1,00	0,67
	LM18	0,125	12960	7	1000	2000	3000	10	3000	1	1	0	0,50	1,00	0,67

Nota: El mejor rendimiento se encuentra en negrita.

6.11 Resultados Generales

Una vez obtenido y evaluado el modelo basado en Anomaly Transformer, se obtiene un avance promedio aproximado de 5 meses y un F1 Score promedio de 0,91, como se muestra en la **Tabla 37**.

Tabla 37. Resultados generales del rendimiento de Anomaly Transformer.

WT	FS	Faults			Faults	Avg. Advance (Months)	F1 Score
		1 Advance (months)	2 Advance (months)	3 Advance (months)			
1	RFE12	3	3	1	3	2.33	0.86
2	RFE12	3	7		2	5.00	0.80
3	RFE12	7	3	7	3	5.67	1.00
4	LM18	8	7		2	7.50	0.8
5	RFE12	3			1	3.00	1.00
6	AD23	3			1	3.00	1.00
7	RFE12	7	6	6	3	6.33	0.86
8	C6	2			1	2.00	0.67
9	AD23	9			1	9.00	1.00
10	C6	6	4		2	5.00	1.00
11	C6	7			1	7.00	1.00
Avg.						5.08	0.91

Nota: En este contexto, "Advance" se refiere al número de meses transcurridos desde la primera alarma hasta la ocurrencia de falla.

Así mismo, en la **Tabla 38** se presentan los resultados para TranAD, con un avance promedio aproximado de 5,29 meses y un F1 Score promedio de 0,94.

Tabla 38. Resultados generales del rendimiento de TranAD.

WT	FS	Faults			Faults	Advance Avg (Months)	F1 Score
		1 Advance (months)	2 Advance (months)	3 Advance (months)			
1	C6	5	3	1	3	3.00	1.00
2	PV18	8	6		2	7.00	1.00
3	PV18	10	3	7	3	6.67	1.00
4	PV18	6	3		2	4.50	1.00
5	RFE12	3			1	3.00	1.00
6	C6	3			1	3.00	0,67
7	LM18	8	2	5	3	5.00	1.00
8	PV18	7			1	7.00	1.00
9	RFE12	10			1	10.00	1.00
10	RFE12	8	6		2	7.00	1.00
11	PV18	2			1	2.00	0.67
Avg.						5.29	0.94

7. Discusión

A continuación, se discuten los resultados para cada uno de los objetivos propuestos:

El objetivo de este proyecto fue aplicar un modelo de predicción capaz de predecir fallas en los aerogeneradores de la CEV. Además, esta investigación ofrece a la comunidad científica y al público en general un proceso integral para la detección de fallas en el módulo IGBT con el uso de datos SCADA en un terreno complejo mediante el uso de técnicas de selección de características para intentar de predecir fallas en mantenimiento predictivo, debido a que, estos aspectos aún no están ampliamente explorados en la literatura científica. Para lograr esta tarea se eligió un modelo de aprendizaje no supervisado, dado que resulta sumamente complicado y costoso etiquetar manualmente los datos antes y durante una falla.

Se realizó un proceso de limpieza y procesamiento de la base de datos SCADA utilizada en este estudio. Se identificaron datos sin etiquetas, valores atípicos y faltantes, y se procedió a filtrar el conjunto de datos para incluir solo aquellos valores correspondientes al funcionamiento normal del aerogenerador, tal como lo hicieron otros estudios previos en detección de anomalías como Santolamazza et al. (2021), Garan et al. (2022) y Rashid y Batunlu (2021). Esto se debe a que los valores nulos y atípicos podrían afectar la precisión de los modelos.

Se realizó la selección de variables según lo recomiendan Ozdemir y Susarla (2018), con diversos métodos como el de RFE, C, PV, AD y LM, en donde RFE es el método que mayor rendimiento obtuvo en la predicción de fallas del módulo IGBT en Anomaly Transformer y el método PV para TranAD. Para un adecuado procesamiento de los datos, se escalaron entre 0 y 1, de acuerdo a lo recomendado por Rashid y Batunlu (2021), Thoresen (2019) y Pertuz, (2022), que destacan que el escalado se realiza para evitar sesgos y facilitar el entrenamiento del modelo. La división de los datos no se realizó como se hace usualmente en modelos de Aprendizaje Automático, debido a que el modelo de Aprendizaje Profundo sin supervisión debe ser entrenado y validado con datos de funcionamiento normal como lo indican Sánchez-Zas et al. (2022), y en los modelos TranAD y AnomalyTransformer de Tuli et al. (2022) y J. Xu et al. (2021) respectivamente.

Se lograron obtener importantes resultados con el modelo Anomaly Transformer, el cual fue capaz de predecir las fallas con un avance aproximado promedio de 5,09 meses y una puntuación F1 Score promedio del 91%. Por otro lado, el modelo TranAD también arrojó resultados prometedores, con un avance aproximado promedio de 5,29 meses y una puntuación F1 Score promedio del 94%. Es importante destacar que la arquitectura Transformador utilizada en este estudio, la cual también es empleada en el modelo DualTransformer de

Chatterjee y Dethlefs (2020) para la detección de fallas en aerogeneradores, aún no ha sido ampliamente utilizada para la predicción de fallas. No obstante, se inició la implementación en este trabajo bajo la hipótesis de que las zonas de alta densidad de anomalías precedieron a la ocurrencia de una falla. Además, es importante mencionar que la arquitectura Transformador empleada en este estudio, puede ser un valioso aporte en otras aplicaciones para la detección de anomalías; además, de los modelos TranAD y Anomaly Transformer, es importante destacar que esta arquitectura ha sido usada con éxito en otros modelos recientes, como AnomalyPTG implementado por Li et al. (2022). Estos modelos han sido aplicados a la detección de anomalías en sistemas de agua (SwAT), en el conjunto de datos del Laboratorio de Ciencias de Marte (MSL), y en la distribución de agua (WADI), obteniendo resultados prometedores. Estos sistemas también monitorean sensores y actuadores de manera similar a los datos SCADA.

Para el etiquetado de alarmas se ha utilizado un enfoque basado en etapas dependiendo de la densidad de anomalías encontradas, con el objetivo de obtener un modelo robusto y reducir los falsos positivos. Para Anomaly Transformer, se tomó una evaluación de una etapa, mientras que para TranAD se empleó un método en dos etapas. Estos métodos basados en etapas se encuentran detalladamente explicados en las secciones de evaluación para cada modelo correspondiente. Otra observación importante incluye que este enfoque de evaluación es similar al sistema Indicador Global de Mahalanobis (GMI) e Indicadores Residuales Locales (LRI), que han sido utilizados en estudios de predicción por Miele et al. (2022), y al método basado en umbral y conteo de anomalías empleados por Chesterman et al. (2021), Campoverde et al. (2022) y Latiffianti et al. (2022).

En general, los resultados presentados tanto para TranAD como para Anomaly Transformer son prometedores y sugieren que existe relación entre las variables del sistema SCADA con las fallas del módulo IGBT. Estos resultados varían según el aerogenerador analizado. Para determinar si los resultados de esta investigación pueden implementarse en la CEV, se debe considerar varios factores como: la estimación del costo de una falla no detectada, el costo de investigar una falsa alarma y el ajuste de parámetros para datos actualizados. Estas estimaciones pueden usarse junto con las métricas de evaluación de esta investigación para determinar si la implementación del sistema es rentable.

8. Conclusiones

- Se realizó el procesamiento de la base de datos del sistema SCADA de la CEV para determinar patrones de comportamiento en las variables, y determinar la integridad de la información. En el procesamiento se encontraron archivos en formato de texto (.txt) y temporales (.tmp), los cuales contenían datos sin etiquetas, valores atípicos y faltantes, estos fueron filtrados para incluir en el estudio únicamente los datos del aerogenerador cuando está en fase de funcionamiento. Además, se analizó la curva de operación de los aerogeneradores y se descubrió que la mayoría de las alarmas se concentran en la zona de operación de máxima potencia. Asimismo, se realizó la selección de variables utilizando los métodos RFE, C, PV, AD y LM y se escaló el conjunto de datos.
- Se implementó TranAD y Anomaly Transformer que son modelos de Aprendizaje Profundo sin supervisión basados en transformadores para la detección de fallas en el módulo IGBT de los aerogeneradores, utilizando como datos de entrada los datos operativos SCADA. Estos modelos son capaces de detectar anomalías en datos de series temporales, por lo que se aplicó un método de evaluación de dos etapas para utilizarlos en la predicción de fallas. Los resultados experimentales muestran que los modelos tienen un buen desempeño en la detección de fallas. Esto demuestra que los modelos de Aprendizaje Profundo sin supervisión basado en transformadores son una herramienta prometedora para la detección temprana de fallas en diferentes componentes de los aerogeneradores de la CEV y pueden contribuir significativamente en las tareas de O&M.
- Se evaluó los dos modelos de predicción y los resultados indican que TranAD obtuvo mejor rendimiento que Anomaly Transformer en todos los aerogeneradores, con un F1 Score promedio del 94% y un avance promedio de aproximadamente 5,29 meses; mientras que Anomaly Transformer obtuvo un F1 Score promedio del 90% y un avance promedio aproximado de 5,08 meses. Esto significa que ambos modelos podrían alertar de fallas incipientes con una precisión y sensibilidad (recall) superior al 90%, y con un avance promedio aproximado de 5 meses.
- La principal contribución de investigación de este proyecto es la implementación de un algoritmo capaz de predecir fallas para aerogeneradores de la CEV, especialmente enfocados en el módulo IGBT basado en datos SCADA. La metodología utilizada para este componente resultó ser aceptable y podrá ser

utilizada en otros componentes. La contribución final de la investigación es que el proceso de predicción de fallas es bastante complejo, y que las diferentes fallas tienen sus propias variables y es muy difícil de determinar el tiempo de ocurrencia de la falla. Sin embargo, la cantidad de alarmas detectadas que no están relacionadas con los reemplazos (FPs) varían según ciertos parámetros del modelo. Además, se destaca un equilibrio entre la precisión del modelo en la predicción de reemplazos y la cantidad de alarmas detectadas que no están relacionadas con ellos. Es importante tener en cuenta que estas últimas alarmas no pueden ser consideradas simplemente como "falsos positivos", ya que es posible que se deban a problemas reales que aún no se han identificado.

9. Recomendaciones

- Para futuros trabajos se sugiere explorar modelos de Aprendizaje Profundo basado en transformadores (TranAD y Anomaly Transformer) en otros componentes mecánicos o eléctricos del aerogenerador, con el objetivo de detectar fallas en etapas tempranas, lo que permitiría optimizar las tareas de mantenimiento preventivo y reducir costos asociados con el cambio de componentes.
- Además de los métodos de selección de variables utilizadas, se podría explorar el uso de técnicas de extracción de características como el análisis de componentes principales y algoritmos de Aprendizaje Automático.
- Se podría considerar el uso de técnicas de aprendizaje por transferencia (Transfer Learning) utilizando datos de aerogeneradores cercanos en los que haya suficiente información normal para entrenar el modelo.
- Para explorar nuevos resultados del modelo TranAD, se recomienda retomar los métodos de selección de características RFE24 y AD23 que quedaron inconclusos debido a limitaciones computacionales. Estos métodos podrían llevar a una mejora en la precisión de las predicciones.
- Para mejorar la precisión en la predicción de fallas, se podría ajustar el método de evaluación de fallas para que tenga en cuenta el tiempo de respuesta necesario para abordar la falla y reducir el riesgo de tiempos de inactividad no planificados y evitar reemplazar el componente con vida útil restante.
- Se podría considerar el uso de una métrica de evaluación basada en el costo de predecir una falla frente al costo de enviar a un técnico a inspeccionar el aerogenerador en el caso de que no haya una falla real, lo que podría ayudar a mejorar el equilibrio entre la detección temprana de fallas y la minimización de falsas alarmas.
- Crear una interfaz gráfica (GUI) donde los usuarios puedan explorar fácilmente los hiperparámetros del modelo para cada aerogenerador.
- Para solicitar el código fuente realizado en este trabajo, que abarca desde el procesamiento de datos, selección de características y la construcción de los modelos predictivos, por favor comuníquese con el director de este proyecto y miembro del CITE, el Ing. Jorge Maldonado, a través de su correo electrónico: jorge.maldonado@unl.edu.ec. También puede contactar al autor, Joel Torres, enviando un correo electrónico a: joel_6197@hotmail.com. Además, en el **Anexo 5** encontrará una guía para replicar los resultados de manera efectiva.

10. Bibliografía

- Alejandro Blanco, M., Marti-Puig, P., Gibert, K., Cusidó, J., & Solé-Casals, J. (2019). A text-mining approach to assess the failure condition of wind turbines using maintenance service history. *Energies*, *12*(10). <https://doi.org/10.3390/en12101982>
- Artigao, E., Martín-Martínez, S., Honrubia-Escribano, A., & Gómez-Lázaro, E. (2018). Wind turbine reliability: A comprehensive review towards effective condition monitoring development. *Applied energy*, *228*, 1569-1583.
- Ayala, M., Maldonado, J., Paccha, E., & Riba, C. (2017). Wind Power Resource Assessment in Complex Terrain: Villonaco Case-study Using Computational Fluid Dynamics Analysis. *Energy Procedia*, *107*, 41-48. <https://doi.org/10.1016/J.EGYPRO.2016.12.127>
- Bashir, D., Montañez, G. D., Sehra, S., Segura, P. S., & Lauw, J. (2020). An information-theoretic perspective on overfitting and underfitting. *Australasian Joint Conference on Artificial Intelligence*, 347-358.
- Blanco-M, A., Gibert, K., Marti-Puig, P., Cusidó, J., & Solé-Casals, J. (2018). Identifying health status of wind turbines by using self organizing maps and interpretation-oriented post-processing tools. *Energies*, *11*(4), 723.
- Campoverde, L., Tutivén, C., Vidal, Y., & Benaláazar-Parra, C. (2022). SCADA Data-Driven Wind Turbine Main Bearing Fault Prognosis Based on Principal Component Analysis. *Journal of Physics: Conference Series*, *2265*(3), 032107.
- Carbajo, R. S., & Bosch, C. (2022). Machine Learning for Wind Turbine Fault Prediction through the Combination of Datasets from Same Type Turbines. *Materials Research Proceedings*, *20*.
- Cevallos-Sierra, J., & Ramos-Martin, J. (2018). Spatial assessment of the potential of renewable energy: The case of Ecuador. *Renewable and Sustainable Energy Reviews*, *81*, 1154-1165. <https://doi.org/10.1016/J.RSER.2017.08.015>
- Chatterjee, J., & Dethlefs, N. (2020). A dual transformer model for intelligent decision support for maintenance of wind turbines. *2020 International Joint Conference on Neural Networks (IJCNN)*, 1-10. <https://doi.org/10.1109/IJCNN48605.2020.9206839>
- Chatterjee, J., & Dethlefs, N. (2021). Scientometric review of artificial intelligence for operations & maintenance of wind turbines: The past, present and future. *Renewable and Sustainable Energy Reviews*, *144*, 111051.
- Chen, J. S., Chen, W. G., Li, J., & Sun, P. (2018). A Generalized Model for Wind Turbine Faulty Condition Detection Using Combination Prediction Approach and Information Entropy. *Journal of Environmental Informatics*, *32*(1).

- Chesterman, X., Verstraeten, T., Daems, P.-J., Nowé, A., & Helsen, J. (2021). Condition monitoring of wind turbines and extraction of healthy training data using an ensemble of advanced statistical anomaly detection models. *Annual Conference of the PHM Society*, 13(1).
- Dai, Z., Yang, Z., Yang, Y., Carbonell, J., Le, Q. v., & Salakhutdinov, R. (2019). Transformer-xl: Attentive language models beyond a fixed-length context. *arXiv preprint arXiv:1901.02860*.
- Dameshghi, A., & Refan, M. H. (2019). Wind turbine gearbox condition monitoring and fault diagnosis based on multi-sensor information fusion of SCADA and DSER-PSO-WRVM method. *International Journal of Modelling and Simulation*, 39(1), 48-72.
- Dehghani, M., Gouws, S., Vinyals, O., Uszkoreit, J., & Kaiser, Ł. (2018). Universal transformers. *arXiv preprint arXiv:1807.03819*.
- Díaz Torres, M. (2020). *Deep learning transformer architecture for predictive business processes monitoring and anomaly detection*.
- Ding, Y. (2019). Data science for wind energy. En *Data Science for Wind Energy* (First). CRC Press. <https://doi.org/10.1201/9780429490972/DATA-SCIENCE-WIND-ENERGY-YU-DING>
- Enrique, T. (2021). La Central Eólica Villonaco. En *Flickr*. <https://flic.kr/p/2kVzC91>
- Eriksson, J. (2020). *Machine Learning for Predictive Maintenance on Wind Turbines: Using SCADA Data and the Apache Hadoop Ecosystem*. <https://www.diva-portal.org/smash/get/diva2:1420733/FULLTEXT01.pdf>
- Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *International conference on machine learning*, 1126-1135.
- Forootan, M. M., Larki, I., Zahedi, R., & Ahmadi, A. (2022). Machine Learning and Deep Learning in Energy Systems: A Review. *Sustainability*, 14(8), 4832. <https://doi.org/https://doi.org/10.3390/su14084832>
- Fuenza Méndez, B. I. (2020). *Mantenimiento predictivo en generadores de Airbus, utilizando aprendizaje profundo*.
- Garan, M., Tidiri, K., & Kovalenko, I. (2022). A data-centric machine learning methodology: application on predictive maintenance of wind turbines. *Energies*, 15(3), 826.
- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. « O'Reilly Media, Inc.»

- Hassid, M., Peng, H., Rotem, D., Kasai, J., Montero, I., Smith, N. A., & Schwartz, R. (2022). *How Much Does Attention Actually Attend? Questioning the Importance of Attention in Pretrained Transformers*. <https://doi.org/10.48550/arxiv.2211.03495>
- Hossain, M. L., Abu-Siada, A., & Muyeen, S. M. (2018). Methods for advanced wind turbine condition monitoring and early diagnosis: A literature review. *Energies*, *11*(5), 1309. <https://doi.org/10.3390/en11051309>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685-695.
- Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2022). Artificial intelligence applications for industry 4.0: A literature-based study. *Journal of Industrial Integration and Management*, *7*(01), 83-111.
- Latiffianti, E., Sheng, S., & Ding, Y. (2022). Wind Turbine Gearbox Failure Detection Through Cumulative Sum of Multivariate Time Series Data. *Frontiers in Energy Research*, *10*(NREL/JA-5000-82585).
- Li, G., Yang, Z., Wan, H., & Li, M. (2022). Anomaly-PTG: A Time Series Data-Anomaly-Detection Transformer Framework in Multiple Scenarios. *Electronics*, *11*(23), 3955.
- Liu, Z., Xiao, C., Zhang, T., & Zhang, X. (2020). Research on fault detection for three types of wind turbine subsystems using machine learning. *Energies*, *13*(2), 460.
- Maldonado-Correa, J., Martín-Martínez, S., Artigao, E., & Gómez-Lázaro, E. (2020). Using SCADA data for wind turbine condition monitoring: A systematic literature review. *Energies*, *13*(12), 3132.
- Maldonado-Correa, J., Valdiviezo-Condolo, M., Viñan-Ludeña, M. S., Samaniego-Ojeda, C., & Rojas-Moncayo, M. (2021). Wind power forecasting for the Villonaco wind farm. *Wind Engineering*, *45*(5), 1145-1159.
- Marino, D. L., Wickramasinghe, C. S., Rieger, C., & Manic, M. (2022). *Self-Supervised and Interpretable Anomaly Detection using Network Transformers*.
- Marti-Puig, P., Blanco-M, A., Cárdenas, J. J., Cusidó, J., & Solé-Casals, J. (2019). Feature selection algorithms for wind turbine failure prediction. *Energies*, *12*(3), 453.
- Mendapara, J. H. (2021). *Investigation of machine learning techniques for wind turbine fault detection using experimental data*.
- Meng, Z., & Xu, X. (2019). A hybrid short-term load forecasting framework with an attention-based encoder–decoder network based on seasonal and trend adjustment. *Energies*, *12*(24), 4612.

- Miele, E. S., Bonacina, F., & Corsini, A. (2022). Deep anomaly detection in horizontal axis wind turbines using graph convolutional autoencoders for multivariate time series. *Energy and AI*, 8, 100145.
- Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing*, 452, 48-62. <https://doi.org/https://doi.org/10.1016/j.neucom.2021.03.091>
- Olabi, A. G., Wilberforce, T., Elsaid, K., Sayed, E. T., Salameh, T., Abdelkareem, M. A., & Baroutaji, A. (2021). A Review on Failure Modes of Wind Turbine Components. *Energies* 2021, Vol. 14, Page 5241, 14(17), 5241. <https://doi.org/10.3390/EN14175241>
- Ozbas, E. E., Aksu, D., Ongen, A., Aydin, M. A., & Ozcan, H. K. (2019). Hydrogen production via biomass gasification, and modeling by supervised machine learning algorithms. *International Journal of Hydrogen Energy*, 44(32), 17260-17268.
- Ozdemir, S., & Susarla, D. (2018). *Feature Engineering Made Easy: Identify unique features from your dataset in order to build powerful machine learning systems*. Packt Publishing Ltd.
- Pertuz, C. P. (2022). *Aprendizaje automático y profundo en python: Una mirada hacia la inteligencia artificial*. Ediciones de la U.
- Ponce-Jara, M. A., Castro, M., Pelaez-Samaniego, M. R., Espinoza-Abad, J. L., & Ruiz, E. (2018). Electricity sector in Ecuador: An overview of the 2007–2017 decade. *Energy Policy*, 113, 513-522. <https://doi.org/10.1016/J.ENPOL.2017.11.036>
- Pumsirirat, A., & Yan, L. (2018). Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine. En *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 9, Número 1). <https://doi.org/10.14569/IJACSA.2018.090103>
- Rashid, H., & Batunlu, C. (2021). Anomaly Detection of Wind Turbine Gearbox based on SCADA Temperature Data using Machine Learning. *renewable energy*, 3, 33.
- Rouhiainen, L. (2018). *Inteligencia artificial*. Madrid: Alienta Editorial.
- Sánchez-Zas, C., Larriva-Novo, X., Villagrà, V. A., Rodrigo, M. S., & Moreno, J. I. (2022). Design and Evaluation of Unsupervised Machine Learning Models for Anomaly Detection in Streaming Cybersecurity Logs. *Mathematics* 2022, Vol. 10, Page 4043, 10(21), 4043. <https://doi.org/10.3390/MATH10214043>
- Sandoval Serrano, L. J. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica; no. 11*.

- Santolamazza, A., Dadi, D., & Introna, V. (2021). A data-mining approach for wind turbine fault detection based on SCADA data analysis using artificial neural networks. *Energies*, *14*(7), 1845.
- Shaw, P., Uszkoreit, J., & Vaswani, A. (2018). Self-attention with relative position representations. *arXiv preprint arXiv:1803.02155*.
- Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*.
- Thoresen, T. H. (2019). *Deep Learning for Fault Prediction in Offshore Wind Turbines*. NTNU.
- Tuli, S., Casale, G., & Jennings, N. R. (2022). *TranAD: Deep transformer networks for anomaly detection in multivariate time series data*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.
- Vilema Lara, P. H. (2022). *Detección de fallos en mantenimiento predictivo utilizando el método de aprendizaje de máquina Random Forest*.
- Vives, J., Quiles, E., & García, E. (2020). AI techniques applied to diagnosis of vibrations failures in wind turbines. *IEEE Latin America Transactions*, *18*(08), 1478-1486.
- Wu, N., Green, B., Ben, X., & O'Banion, S. (2020). Deep transformer models for time series forecasting: The influenza prevalence case. *arXiv preprint arXiv:2001.08317*.
- Wu, S., Xiao, X., Ding, Q., Zhao, P., Wei, Y., & Huang, J. (2020). Adversarial sparse transformer for time series forecasting. *Advances in neural information processing systems*, *33*, 17105-17115.
- Xinjiang Goldwind Science & Technology Co., Ltd. (2011). *GW70/1500 Wind Turbine Technical Description*.
- Xu, J., Wu, H., Wang, J., & Long, M. (2021). *Anomaly transformer: Time series anomaly detection with association discrepancy*.
- Xu, L., Xu, K., Qin, Y., Li, Y., Huang, X., Lin, Z., Ye, N., & Ji, X. (2022). TGAN-AD: Transformer-Based GAN for Anomaly Detection of Time Series Data. *Applied Sciences*, *12*(16), 8085. <https://doi.org/10.3390/app12168085>
- Yang, H. H., Huang, M. L., Lai, C. M., & Jin, J. R. (2018). An approach combining data mining and control charts-based model for fault detection in wind turbines. *Renewable Energy*, *115*, 808-816. <https://doi.org/10.1016/J.RENENE.2017.09.003>

Zhang, H., Zhang, L., & Jiang, Y. (2019). Overfitting and underfitting analysis for deep learning based end-to-end communication systems. *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, 1-6.

11. Anexos

Anexo 1. Variables SCADA

En la **Tabla 39** se encuentra identificado las variables del sistema SCADA por ID y su descripción.

Tabla 39. Anexo de descripción de variables SCADA.

ID	Variable	Unidad	Descripción
1	wind_speed_avg	m/s	Velocidad de viento promedio. Se mide con el anemómetro de góndola
2	wind_speed_max	m/s	Velocidad de viento máxima. Se mide con el anemómetro de góndola
3	wind_speed_min	m/s	Velocidad de viento mínima. Se mide con el anemómetro de góndola
4	grid_active_power_avg	kW	Potencia activa promedio
5	grid_active_power_max	kW	Potencia activa máxima
6	grid_active_power_min	kW	Potencia activa mínima
7	generator_capacitors_temperature_max	°C	Temperatura máxima del capacitor del generador
8	converter_reactive_power_avg	kVAR	Potencia reactiva promedio en el convertidor
9	converter_reactive_power_max	kVAR	Potencia reactiva máxima en el convertidor
10	converter_reactive_power_min	kVAR	Potencia reactiva mínima del convertidor
11	generator_speed_avg	rpm	Velocidad rotación promedio del generador. (rango 0-20 rpm)
12	generator_speed_max	rpm	Velocidad rotación máxima del generador.
13	generator_speed_min	rpm	Velocidad rotación mínima del generador.
14	ambient_temperature_avg	°C	Temperatura ambiente promedio
15	ambient_temperature_max	°C	Temperatura ambiente máxima
16	ambient_temperature_min	°C	Temperatura ambiente mínima
17	blade_angle_avg	°(degree)	Ángulo promedio de la pala (rango -5 - 95°)
18	blade_angle_max	°(degree)	Ángulo máximo de la pala
19	blade_angle_min	°(degree)	Ángulo mínimo de la pala
20	nacelle_position_avg	°(degree)	Posición del nacelle (controlado por yaw) (rango 0 - 900)
21	grid_U1_avg	V	Tensión promedio 1
22	grid_U2_avg	V	Tensión promedio 2
23	grid_U3_avg	V	Tensión promedio 3
24	grid_I1_avg	A	Corriente promedio 1
25	grid_I2_avg	A	Corriente promedio 2
26	grid_I3_avg	A	Corriente promedio 3
27	acceleration_nacelle_max	m/s ²	Aceleración de la góndola. (rango -0.5g - 0.5g) medición en X y Y
28	winding_temperature_max	°C	Temperatura máxima del bobinado del generador
29	topbox_temperature_max	°C	Temperatura máxima en la caja del sistema de control de la góndola
30	nacelle_temperature_max	°C	Temperatura máxima dentro de la góndola
31	igbt_temperature_max	°C	Temperatura máxima del IGBT
32	dc_link_capacitors_temperature_max	°C	Temperatura máxima de capacitores dc link
33	ac_inductor_temperature_max	°C	Temperatura máxima de ac inductor
34	rectifier_temperature_max	°C	Temperatura máxima del rectificador
35	chopper_igbt_temperature_max	°C	Temperatura máxima en chopper IGBT
36	dc_inductor_temperature_max	°C	Temperatura máxima en dc inductor
37	step_up_igbt_temperature_max	°C	Temperatura máxima de step-up IGBT
38	pitch_motor_temperature_1_max	°C	Temperatura máxima de motor de pich 1
39	pitch_motor_temperature_2_max	°C	Temperatura máxima de motor de pich 2
40	pitch_motor_temperature_3_max	°C	Temperatura máxima de motor de pich 3
41	pitch_capacitor_temperature_1_max	°C	Temperatura máxima de capacitor de pich 1
42	pitch_capacitor_temperature_2_max	°C	Temperatura máxima de capacitor de pich 2
43	pitch_capacitor_temperature_3_max	°C	Temperatura máxima de capacitor de pich 3
44	pitch_cabinet_temperature_1_max	°C	Temperatura máxima de gabinete de pich 1
45	pitch_cabinet_temperature_2_max	°C	Temperatura máxima de gabinete de pich 2
46	pitch_cabinet_temperature_3_max	°C	Temperatura máxima de gabinete de pich 3
47	pitch_converter_temperature_1_max	°C	Temperatura máxima de convertidor de pich 1
48	pitch_converter_temperature_2_max	°C	Temperatura máxima de convertidor de pich 2
49	pitch_converter_temperature_3_max	°C	Temperatura máxima de convertidor de pich 3
50	pitch_power_supply_temperature_1_max	°C	Temperatura máxima de power supply de pich 1
51	pitch_power_supply_temperature_2_max	°C	Temperatura máxima de power supply de pich 2

ID	Variable	Unidad	Descripción
52	pitch_power_supply_temperature_3_max	°C	Temperatura máxima de power supply de pitch 3
53	operation_mode	número	N/A
54	data_energy_yield	kW/h	Energía generada acumulada
55	data_consumed_energy_yield	kW/h	Energía consumida
56	data_Power_on_time	horas	N/A
57	data_WTG_ok_time	horas	N/A
58	data_error_time	horas	N/A
59	data_environment_ok_time	horas	N/A
60	data_environment_not_ok_time	horas	N/A
61	data_service_time	horas	N/A
62	data_grid_control_standstill_time	horas	N/A
63	data_power_production_time	horas	N/A
64	today_data_energy_yield	kW/h	Energía generada
65	yesterday_data_energy_yield	kW/h	Energía generada acumulada
66	two_day_before_data_energy_yield	kW/h	Energía generada acumulada
67	three_days_before_data_energy_yield	kW/h	Energía generada acumulada

Anexo 2. Códigos de falla y descripción de alarmas SCADA

Tabla 40. Descripción de códigos de falla del sistema SCADA

Fault Code	Fault description	Value
1	Error acceleration nacelle limit	5248
435	Error_converter not ready	1091
442	Error_converter IGBT ok signal loss	199
36	Voltage low	193
39	Current unsymmetry	191
30	S/N	166
496	Error converter feedback	152
503	Error IGBT_ok	151
451	Error_converter DC link voltage high	144
506	Error converter grid IGBT	125
598	error_yaw_acceleration_nacelle_limit	110
497	Error converter trigger	103
508	Converter grid IGBT over-current _converter	95
531	Error converter step-up DC voltage max	92
762	S/N	92
481	Error IGBT fan 1 feedback	82
482	Error IGBT fan 1 operation	82
485	Error IGBT fan 2 feedback	82
486	Error IGBT fan 2 operation	82
422	Error_converter capacitor cooling fan feedback...	78
504	Error converter step-up IGBT	76
505	Error brake IGBT	71
96	Error safety system rotor lock	64
507	Converter step-up over-current _converter	63
38	High current	62
519	Converter grid IGBT temperature unsymmetry (ne...	60
34	Over-high voltage	57
447	Error_converter grid side IGBT over current	52
49	Error grid reactive power max	50
597	error_converter_main_contactor	50
509	Error converter grid voltage max_converter	43
458	Error_converter grid side IGBT temperature uns...	42
35	Voltage high	38
60	Error generator speed critical speed 1	32
470	Error_converter step up DC link voltage high	30
510	Brake system over-current _converter	27
448	Error_converter grid side voltage high	25
42	Error grid active power limit max	24
446	Error_converter step up circuit over current	23
512	Error DC voltage max _converter	22
2	Error acceleration nacelle limit offset	18

Fault Code	Fault description	Value
439	Error_generator breaker	18
130	Error wind anemometer	17
62	Error generator speed mode maintenance	15
433	Error_converter 2# IGBT cabinet temperature	12
420	Error_converter IGBT cooling 1# fan feedback loss	10
523	Converter step-up IGBT temperature unsymmetry ...	10
76	S/N	9
596	error_converter_precharge_not_finish	8
50	Error_grid_reactive power low	5
484	Error capacitor fan operation	5
483	Error capacitor fan feedback	5
452	Error_converter grid side L1a IGBT temperature...	5
220	Error pitch 3 speed limit	5
61	Error generator speed critical speed 2	2
449	Error_converter chopper over current	2
64	Error_Gen low temperature	2
444	Error_converter chopper IGBT fault	2
424	Error_converter IGBT cooling 2# fan feedback loss	2
185	Error_2# pitch motor over temperature	2
454	Error_converter grid side L2a IGBT temperature...	1
186	Error_3# pitch motor over temperature	1
453	Error_converter grid side L1b IGBT temperature...	1
218	Error pitch 1 speed limit	1
219	Error pitch 2 speed limit	1
432	Error_converter 1# IGBT cabinet temperature	1

Anexo 3. Resultados detallados para el primer modelo Anomaly Transformer

I. Detección de fallas en WT2 mediante Anomaly Transformer (detalladamente)

A. Anomaly Transformer utilizando el método RFE1 en WT2

En la **Figura 52** se observa que el puntaje de anomalías supera el umbral de primera etapa donde se detectaron las 2 fallas registradas del módulo IGBT. Sin embargo, también se pueden ver una advertencia detectada donde no hubo falla a inicios de 2019 y también a finales de 2021 dando dos FPs.

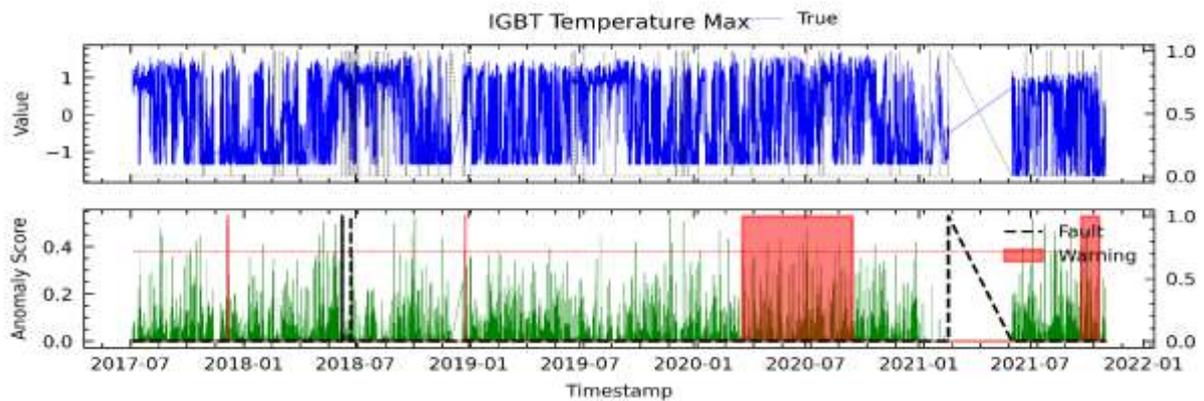


Figura 52. Resultados para RFE1 en el análisis fallas en el módulo IGBT mediante el modelo AT para WT2.

B. Anomaly Transformer utilizando el método RFE12 en WT2

En la **Figura 53** se presenta el resultado del análisis de la falla del módulo IGBT con las variables del método RFE12, dando como resultado las 2 fallas advertidas con un avance promedio aproximado de 4 meses y un FP a finales de 2021.

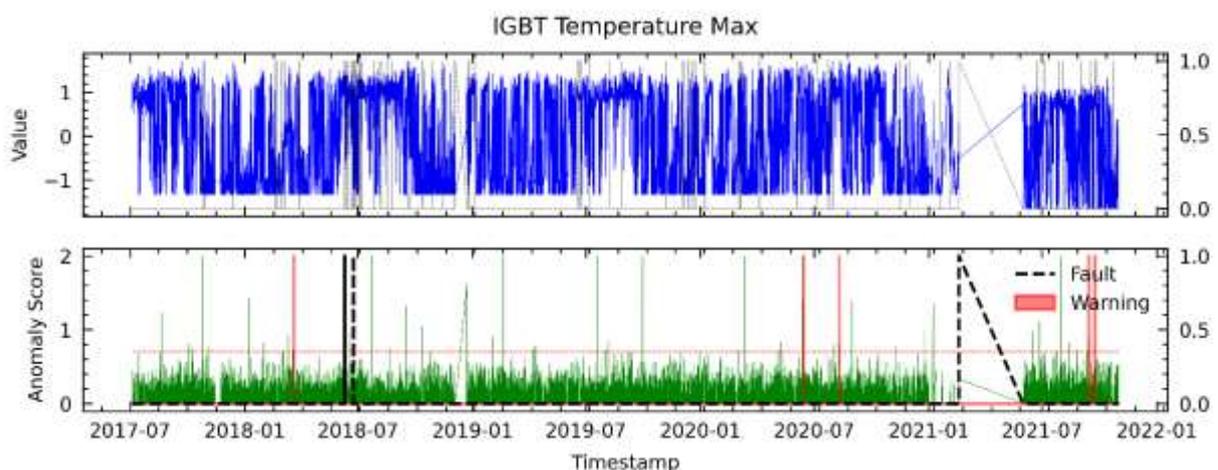


Figura 53. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.

C. Anomaly Transformer utilizando el método RFE24 en WT2

En la **Figura 54** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que se detectó las 2 fallas, pero en cambio se registra 2 FPs prolongados.

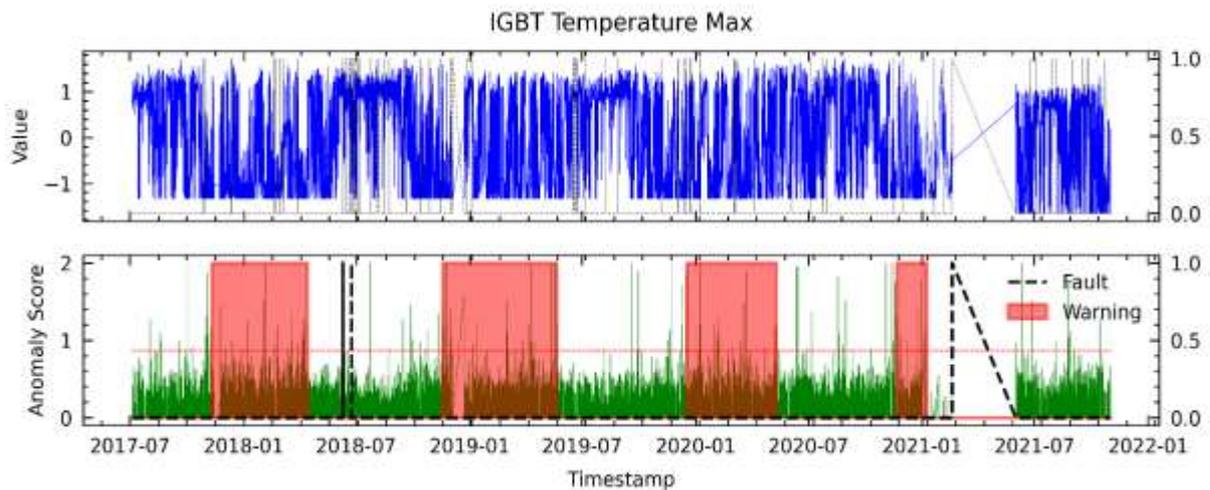


Figura 54. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.

D. Anomaly Transformer utilizando el método C6 en WT2

En el método de selección de variables por correlación se eligieron 6 variables como entrada al modelo, resultando en la **Figura 55**. Se puede observar que el modelo logró detectar las dos fallas del módulo IGBT con hasta seis meses promedio aproximado de anticipación, pero también reportó un FP.

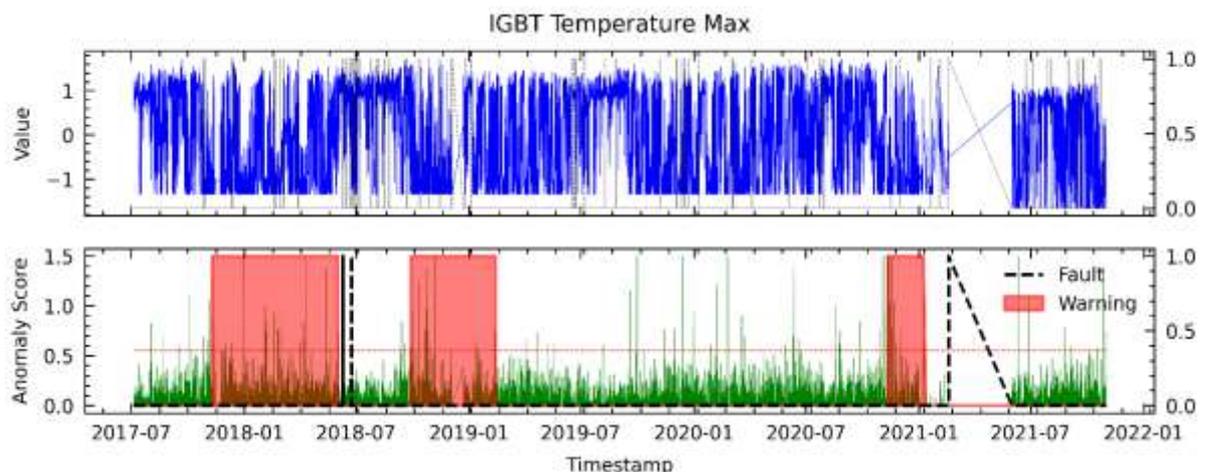


Figura 55. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.

E. Anomaly Transformer utilizando el método PV18 en WT2

La **Figura 56** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió las dos fallas del módulo IGBT, pero también hubo dos alarmas que no conllevaron a una falla del IGBT.

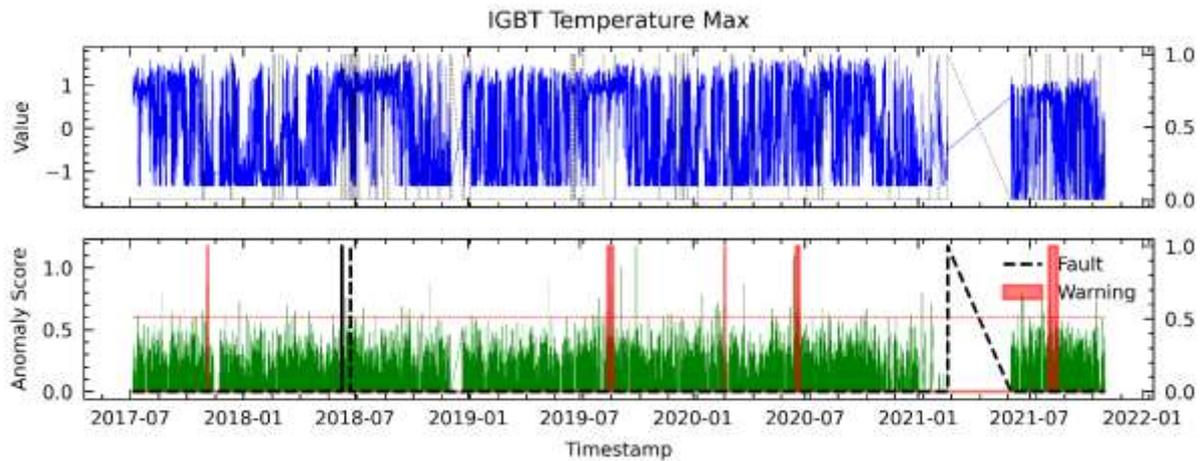


Figura 56. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.

F. Anomaly Transformer utilizando el método AD23 en WT2

En la **Figura 57** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados no son satisfactorios, ya que, aunque es capaz de detectar las dos fallas en el módulo IGBT con hasta tres meses de anticipación aproximadamente, también se han producido tres falsas alarmas.

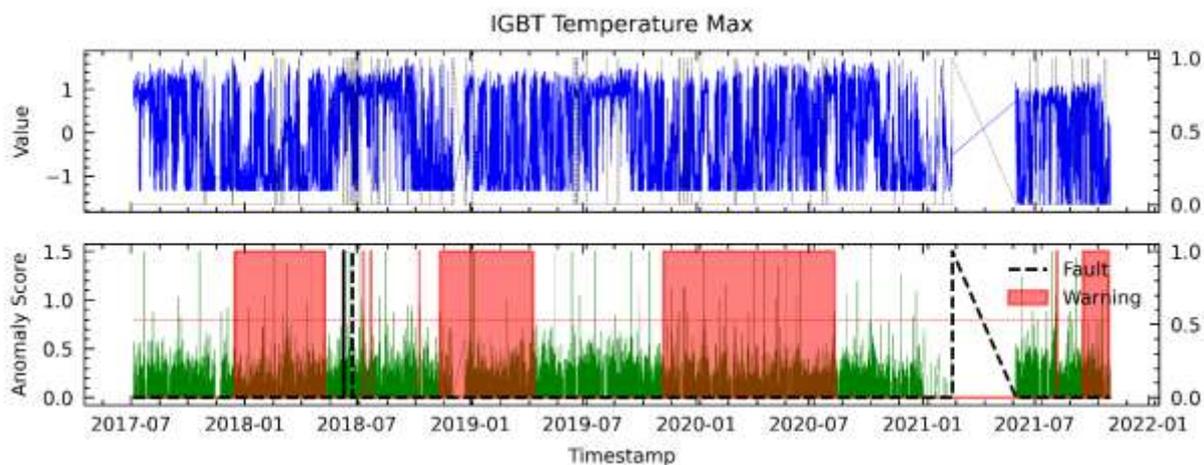


Figura 57. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.

G. Anomaly Transformer utilizando el método LM18 en WT2

En la **Figura 58** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que se lograron detectar las fallas en el módulo IGBT con hasta 3 meses de anticipación aproximadamente, pero también se presentaron 2 FPs.

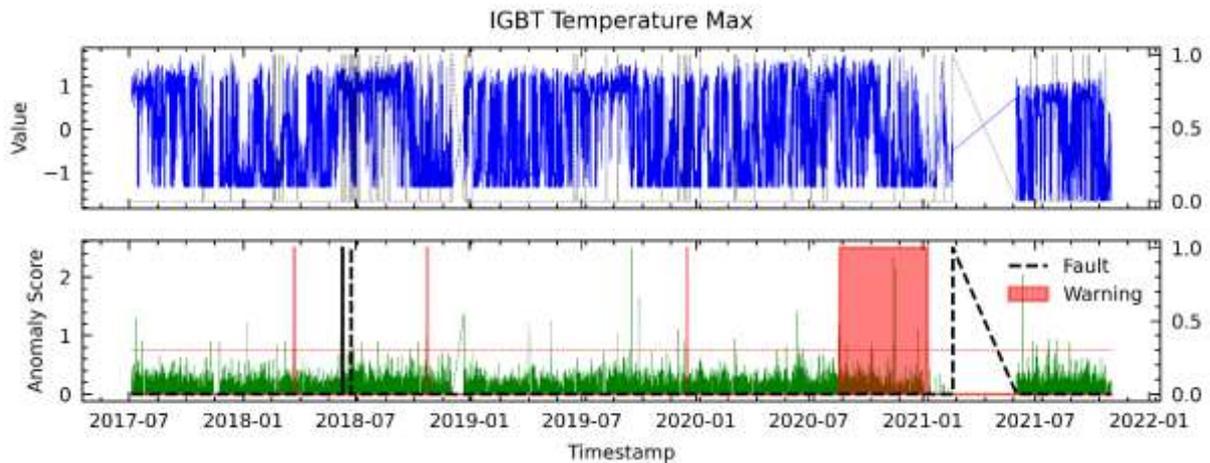


Figura 58. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT2.

II. Detección de fallas en WT3 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT3

En la **Figura 59** se observa que el puntaje de anomalías supera el umbral de primera etapa incluso hasta con un avance promedio de 4 meses aproximadamente antes de la falla del componente IGBT, donde se detectaron las tres fallas registradas. Sin embargo, también se pueden ver dos advertencias del modelo donde no se encuentra registro en junio de 2018 y finales de 2021.

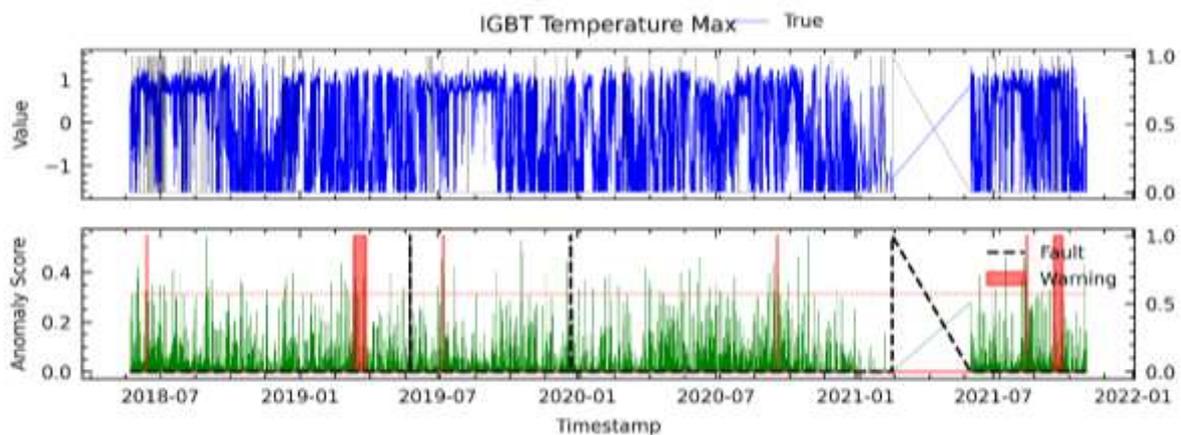


Figura 59. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

B. Anomaly Transformer utilizando el método RFE12 en WT3

En la **Figura 60** se presenta el resultado del análisis de la falla del módulo IGBT con el método de RFE12, dado como resultado las tres fallas detectadas con un avance promedio aproximado de 6 meses.

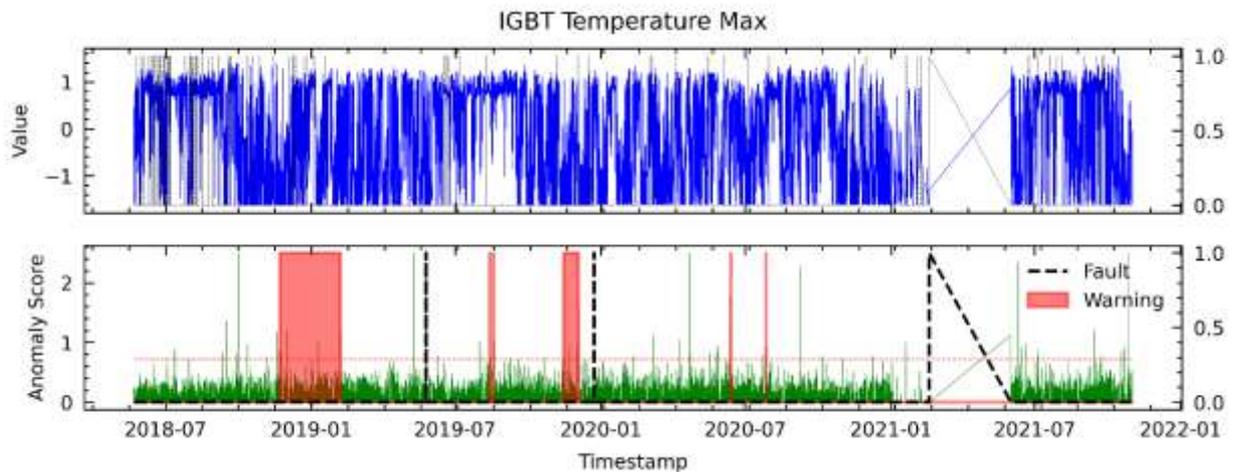


Figura 60. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

C. Anomaly Transformer utilizando el método RFE24 en WT3

En la **Figura 61** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que se detectó las 3 fallas con un avance aproximado de 7 meses con anticipación, pero en cambio se registra 2 FPs.

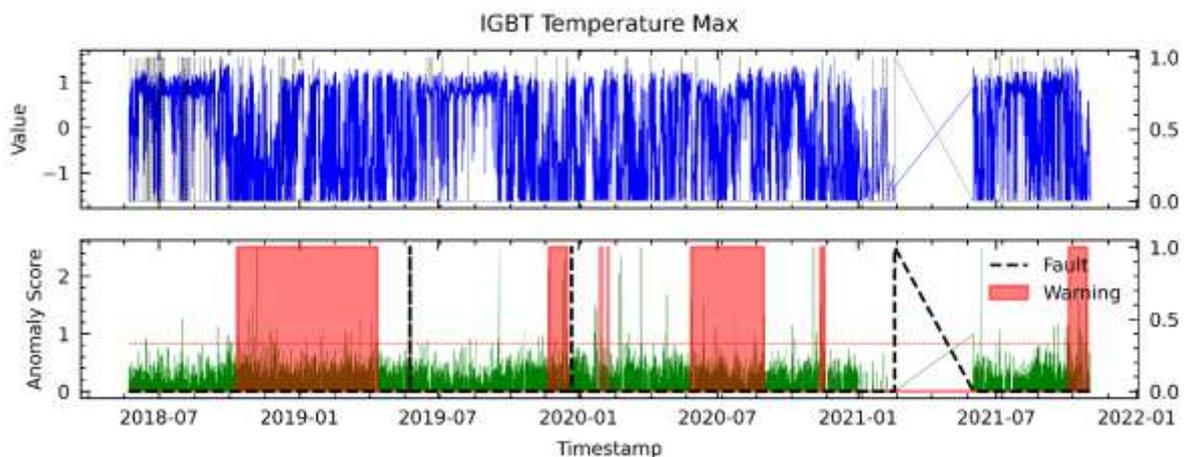


Figura 61. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

D. Anomaly Transformer utilizando el método C6 en WT3

En el método de selección de variables por correlación se eligieron variables como entrada al modelo, resultando en la **Figura 62**. Se puede observar que el modelo logró detectar las tres fallas del módulo IGBT con hasta seis meses de antelación, pero también reportó un FP.

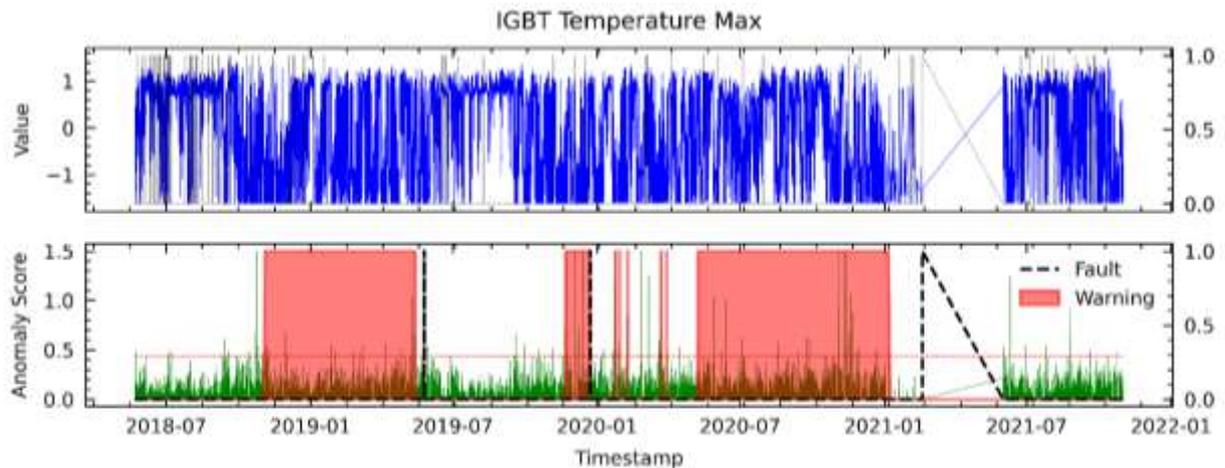


Figura 62. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

E. Anomaly Transformer utilizando el método PV18 en WT3

La **Figura 63** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió las tres fallas del módulo IGBT con un avance aproximado de 4 meses.

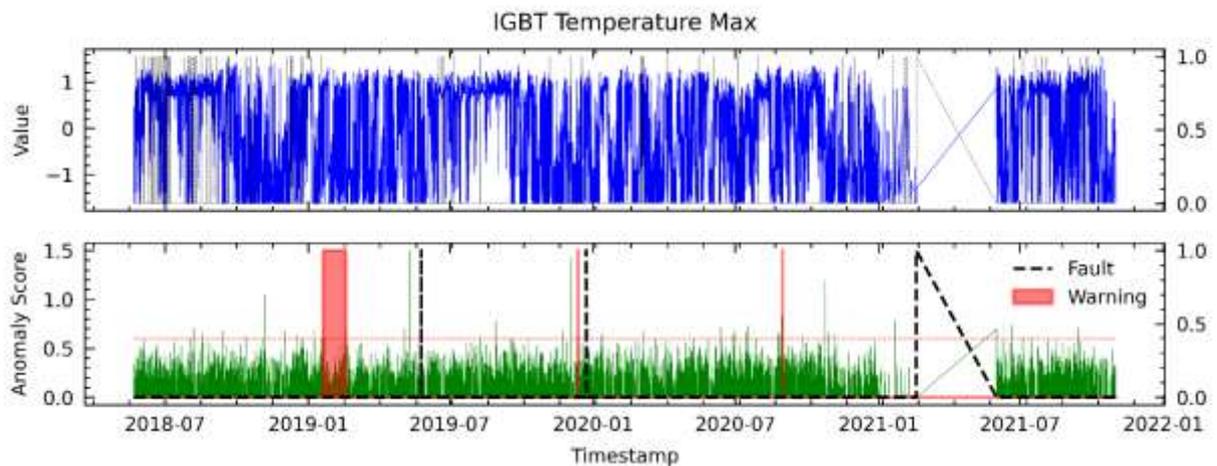


Figura 63. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

F. Anomaly Transformer utilizando el método AD23 en WT3

En la **Figura 64** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados no son satisfactorios, ya que, aunque es capaz de detectar las dos fallas en el módulo IGBT con un promedio aproximado de 5 meses de anticipación, también se han producido un FP.

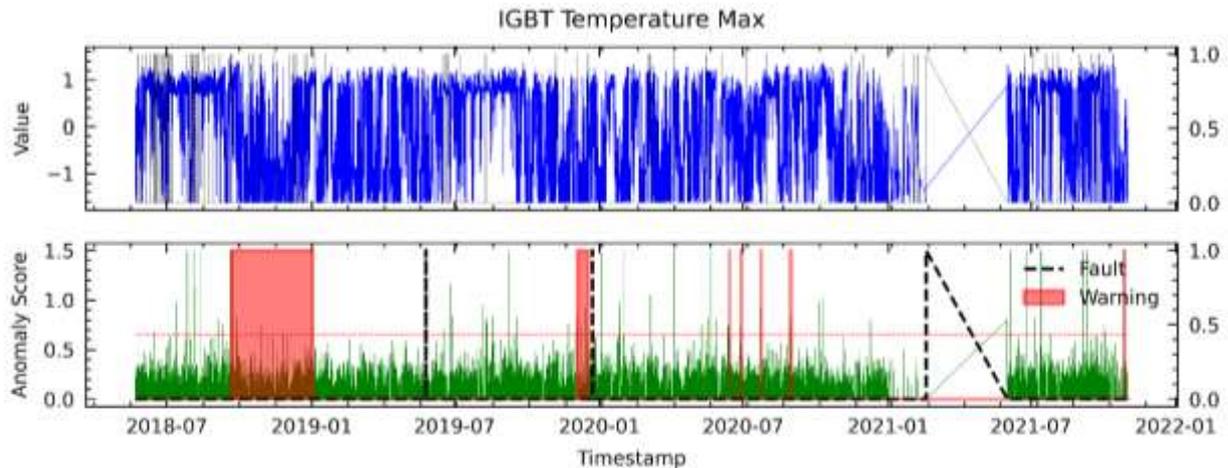


Figura 64. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

G. Anomaly Transformer utilizando el método LM18 en WT3

En la **Figura 65** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que se lograron detectar las fallas en el módulo IGBT con hasta 6 meses de anticipación aproximadamente, pero también se presentaron 3 falsos positivos.

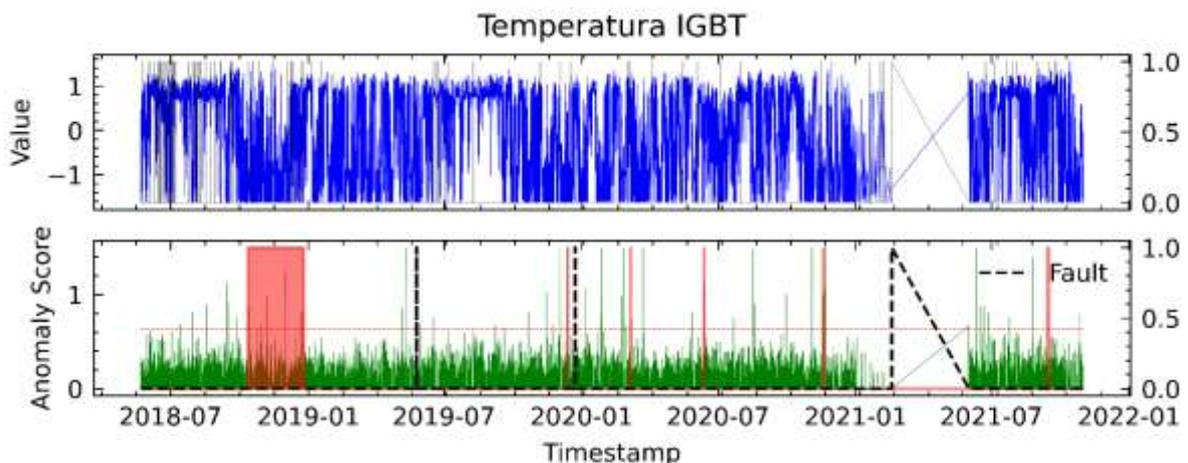


Figura 65. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT3.

III. Detección de fallas en WT4 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT4

En la **Figura 66** se observa que el puntaje de anomalías supera el umbral de primera etapa con un avance aproximado de 6 meses las dos fallas registradas del módulo IGBT. Sin embargo, también se pueden ver una advertencia del modelo donde no se encuentra registro de fallo del módulo IGBT en junio de 2018.

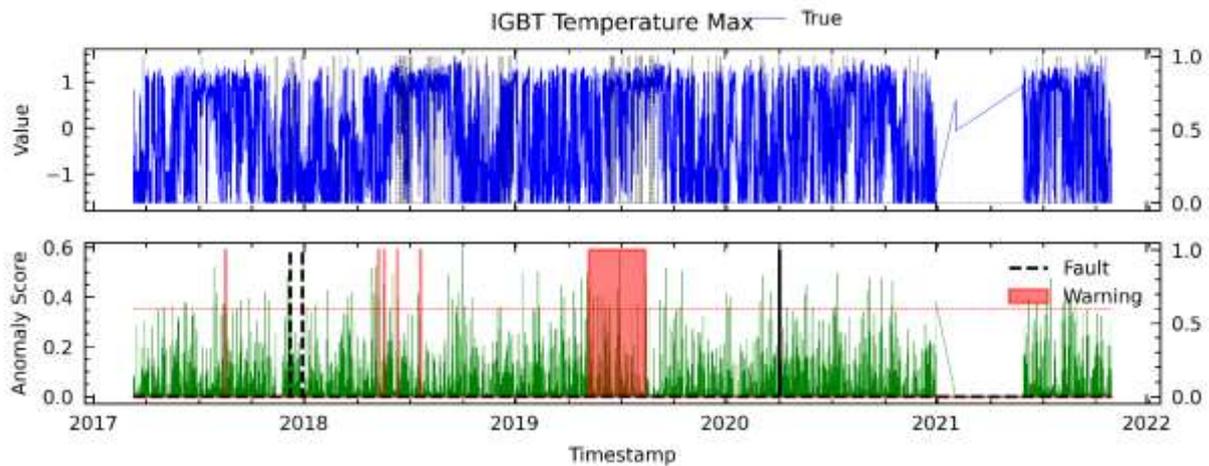


Figura 66. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

B. Anomaly Transformer utilizando el método RFE12 en WT4

En la **Figura 67** se presenta el resultado del análisis de la falla del módulo IGBT con el método RFE12, donde se observa que el modelo solo pudo detectar una de las fallas registradas del IGBT. La advertencia del modelo correspondió hasta 6 meses aproximadamente antes de la falla registrada.

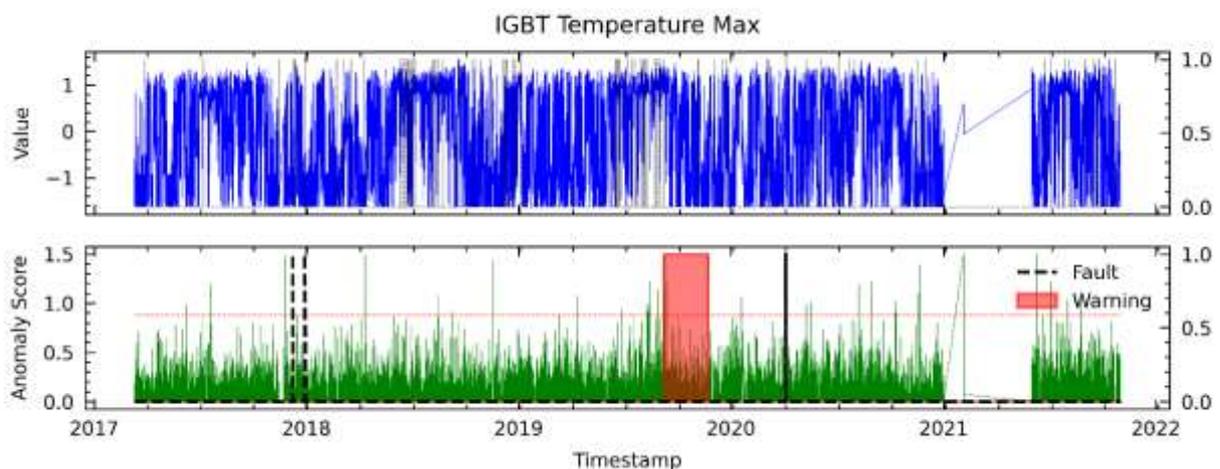


Figura 67. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

C. Anomaly Transformer utilizando el método RFE24 en WT4

En la **Figura 68** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que se detectó las dos fallas, pero en cambio se registra 2 FPs.

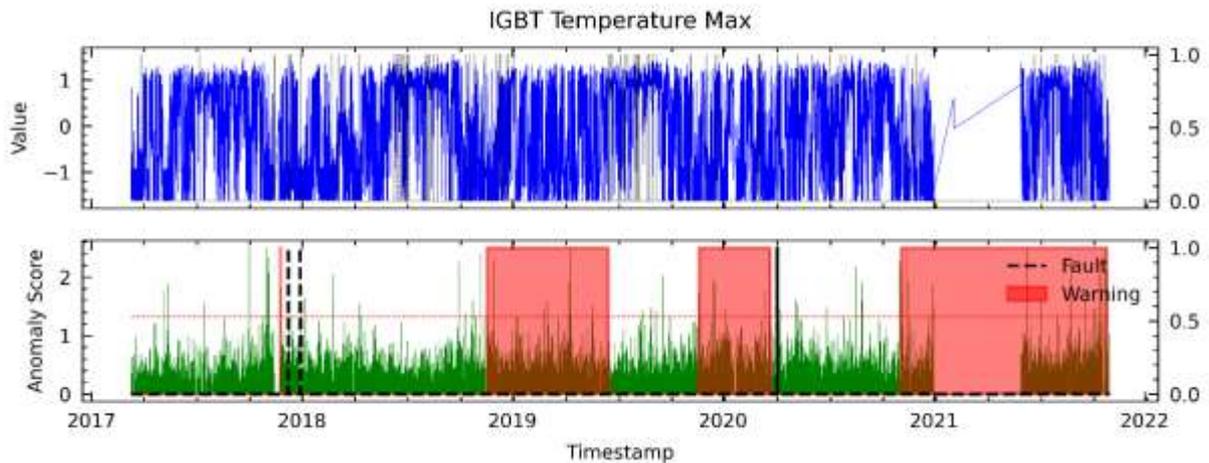


Figura 68. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

D. Anomaly Transformer utilizando el método C6 en WT4

En el método de selección de variables por correlación resulta en la **Figura 69**, donde se puede observar que el modelo logró detectar las dos fallas en el módulo IGBT con un avance aproximado de 6 meses de anticipación, pero también reportó dos FPs.

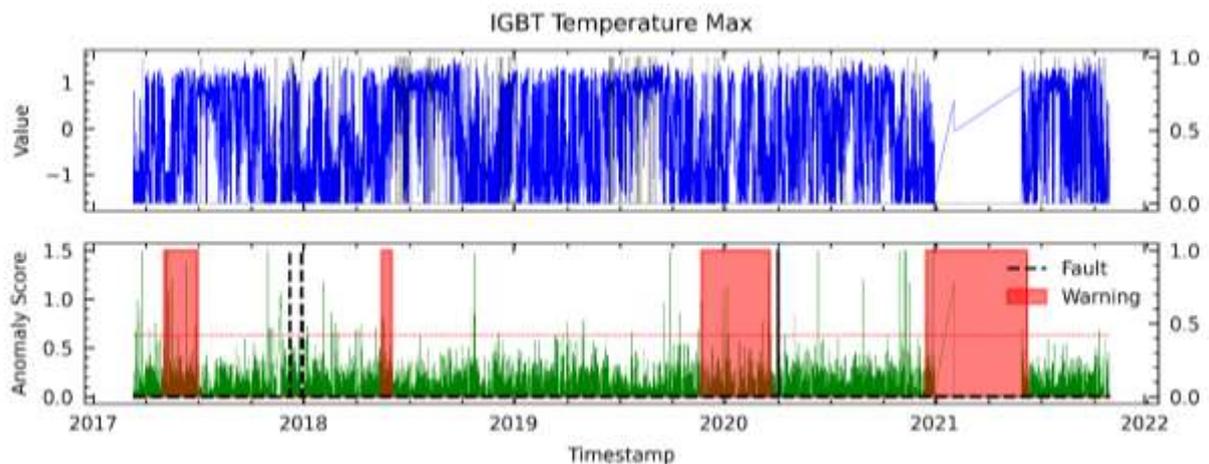


Figura 69. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

E. Anomaly Transformer utilizando el método PV18 en WT4

En la **Figura 70** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió una de las fallas del módulo IGBT y hubo dos FPs.

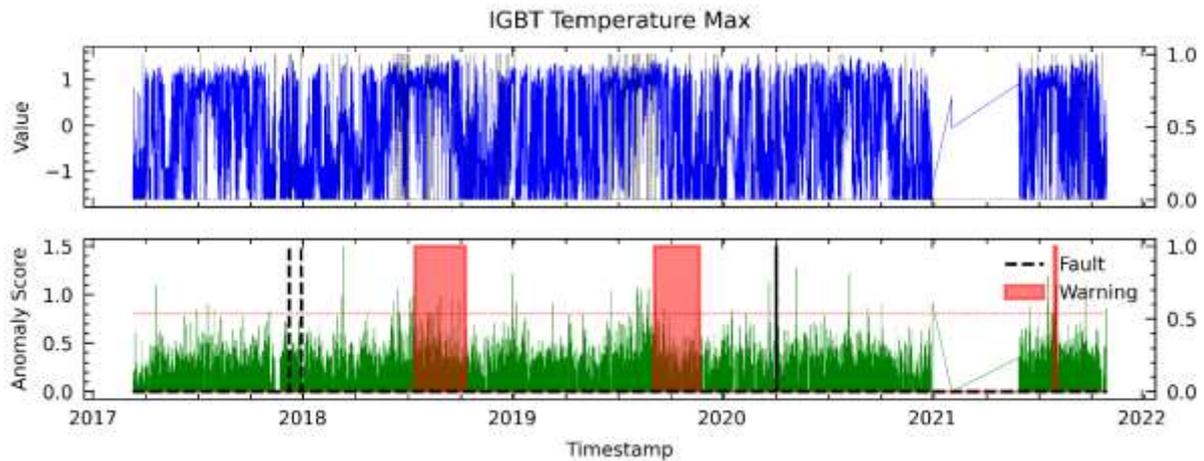


Figura 70. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

F. Anomaly Transformer utilizando el método AD23 en WT4

En la **Figura 71** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados no son satisfactorios, ya que, aunque no es capaz de detectar las dos fallas en el módulo IGBT y registra 2 FPs.

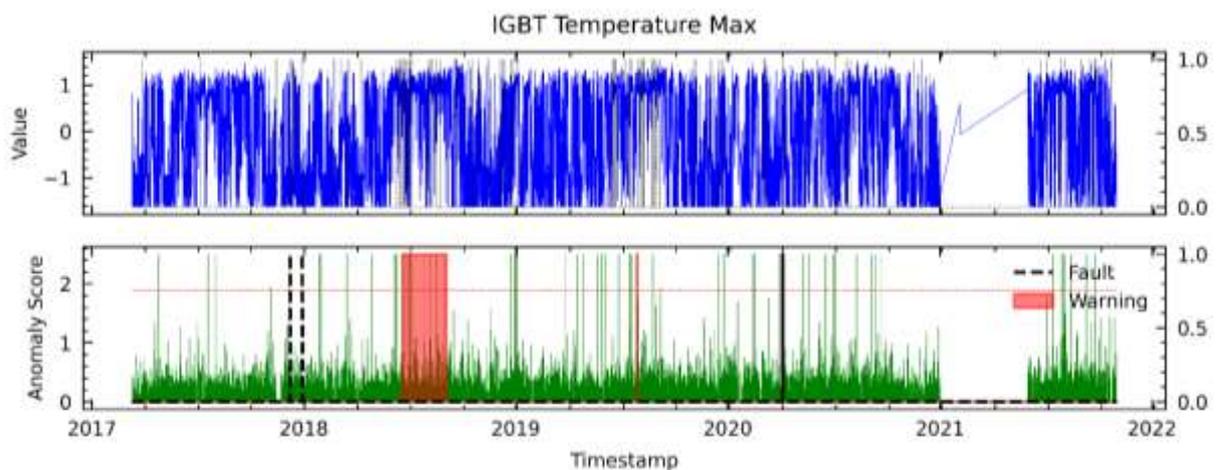


Figura 71. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

G. Anomaly Transformer utilizando el método LM18 en WT4

En la **Figura 72** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que se lograron detectar las dos fallas en el módulo IGBT con un avance promedio aproximado de 7 meses de anticipación, pero también se presentó 1 FP.

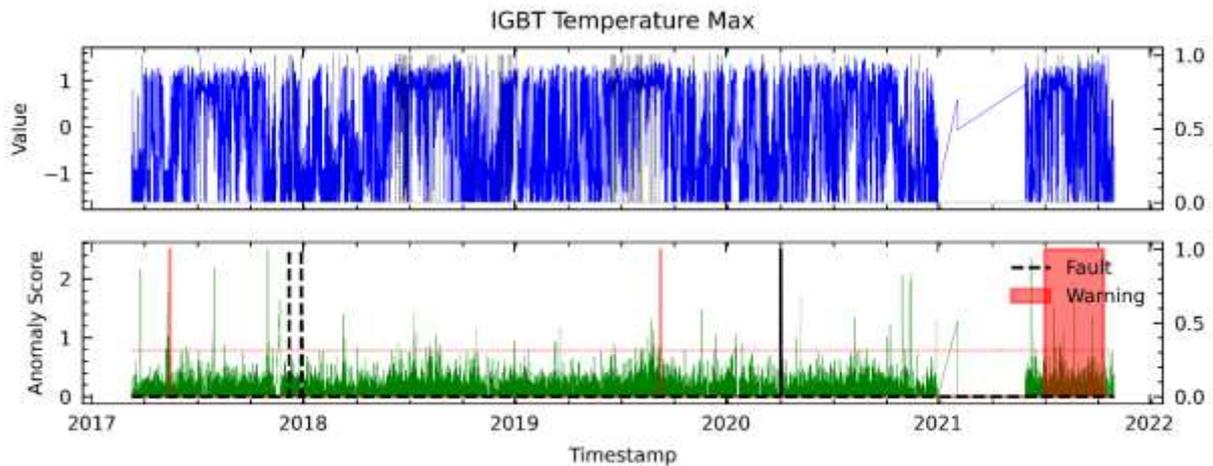


Figura 72. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

IV. Detección de fallas en WT5 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT5

En la **Figura 73** se observa que el puntaje de anomalías supera el umbral de primera etapa de un mes aproximado de anticipación de la falla. Sin embargo, también se pueden ver tres advertencias del modelo donde no se encuentra registro de fallo siendo FPs.

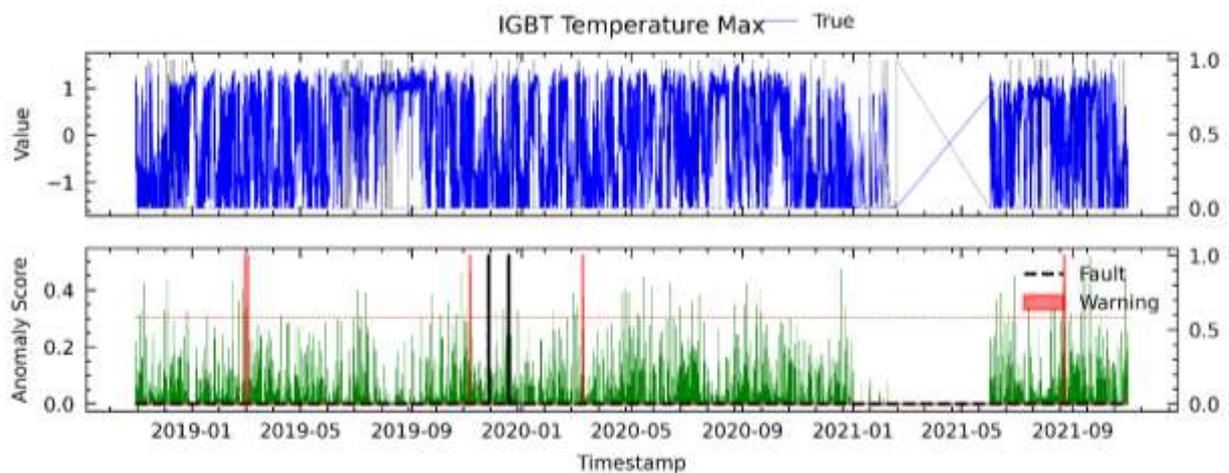


Figura 73. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.

B. Anomaly Transformer utilizando el método RFE12 en WT5

En la **Figura 74** se presenta el resultado del análisis de la falla del módulo IGBT con el método RFE12, donde se observa que el modelo advirtió la falla con un avance aproximado de 3 meses de anticipación.

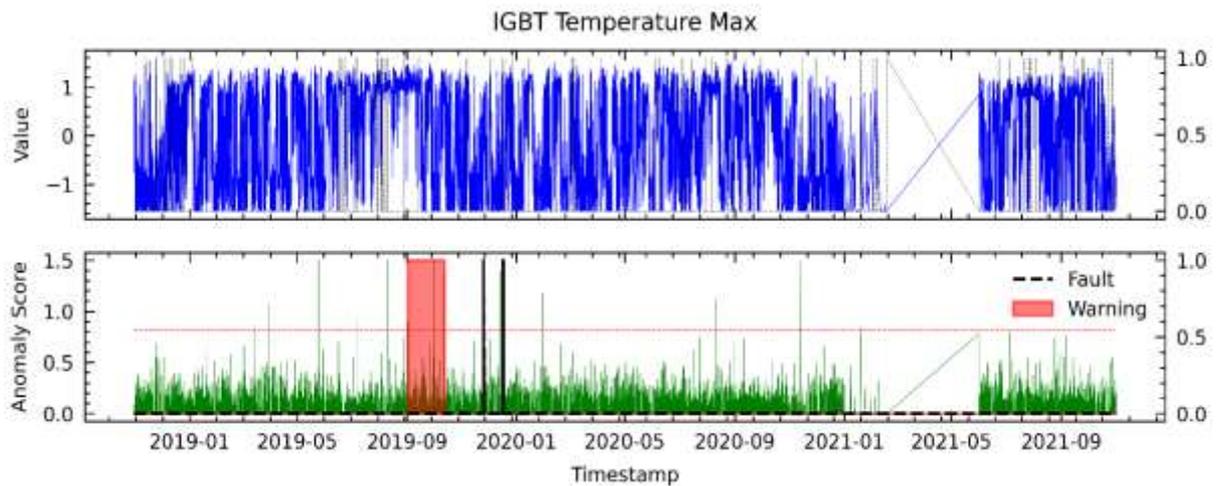


Figura 74. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.

C. Anomaly Transformer utilizando el método RFE24 en WT5

En la **Figura 75** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RF24. Se observa que no se detectó la falla, pero en cambio se registra 1 FP que prevé una posible segunda falla de algún componente desconocido, se tomó esta opinión debido a que los datos SCADA no se encuentran por aproximadamente 2 meses seguidos.

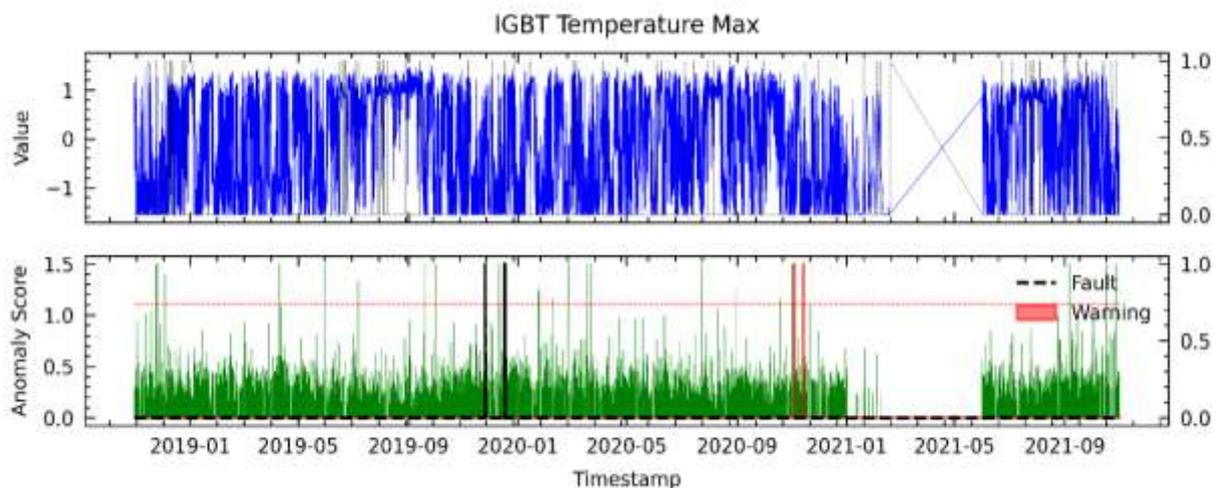


Figura 75. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.

D. Anomaly Transformer utilizando el método C6 en WT5

En el método de selección de variables por correlación se eligieron variables por el método C6 como entrada al modelo, resultando en la **Figura 76**. Se puede observar que el modelo logró detectar la falla en el módulo IGBT con un avance aproximado de 6 meses de anticipación, pero también reportó tres FPs.

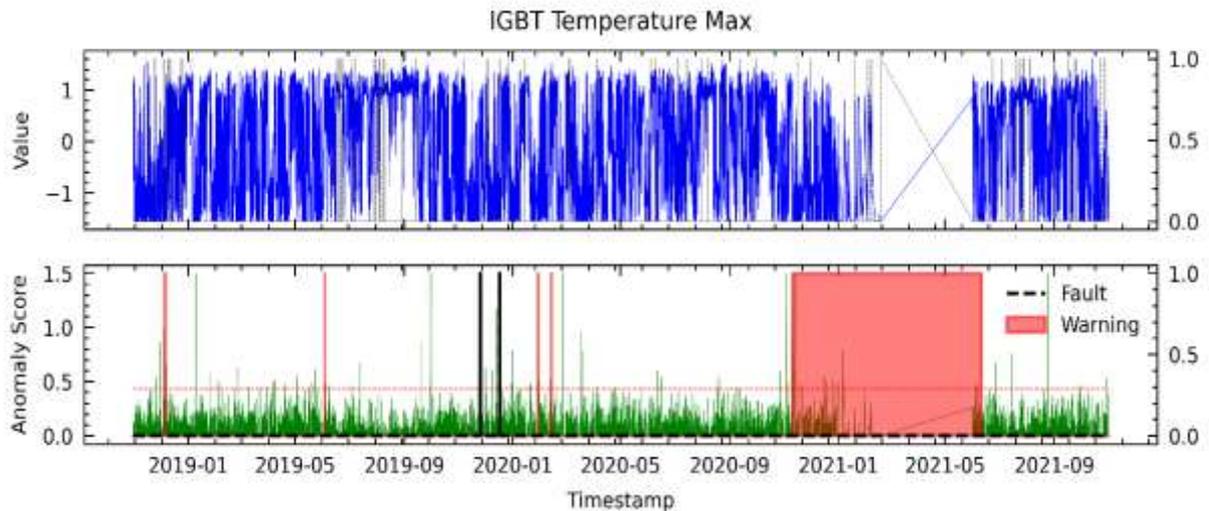


Figura 76. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT4.

E. Anomaly Transformer utilizando el método PV18 en WT5

La **Figura 77** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió las fallas del módulo IGBT en un avance aproximado de dos semanas y también registró dos falsos positivos.

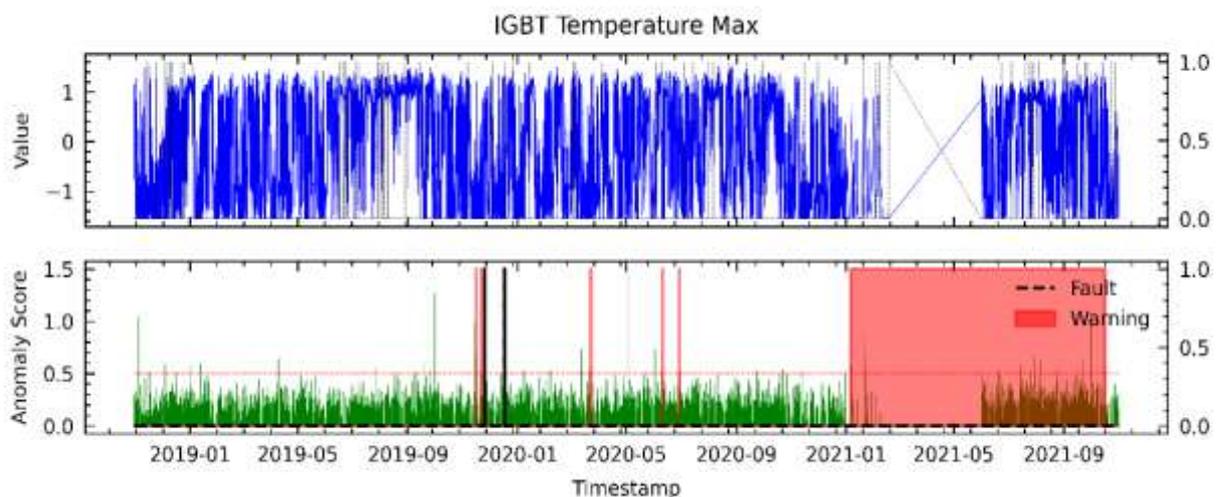


Figura 77. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.

F. Anomaly Transformer utilizando el método AD23 en WT5

En la **Figura 78** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados son satisfactorios, ya que, aunque es capaz de detectar una de la falla en el módulo IGBT con un avance aproximado de 3 meses.

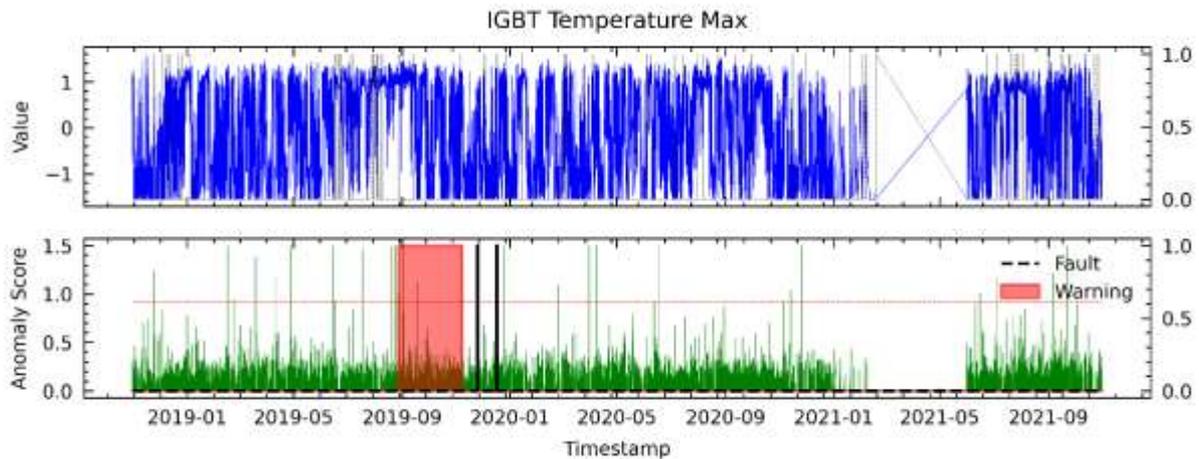


Figura 78. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.

G. Anomaly Transformer utilizando el método LM18 en WT5

En la **Figura 79** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que no se logró detectar la falla en el módulo IGBT y registró un FP.

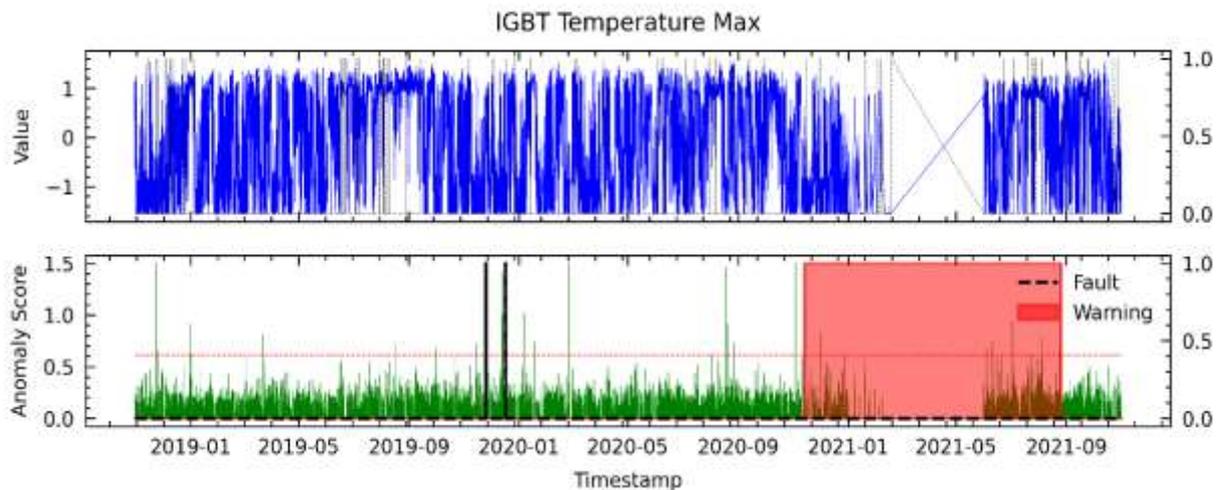


Figura 79. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT5.

V. Detección de fallas en WT6 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT6

En la **Figura 80** se observa que el puntaje de anomalías supera el umbral de primera etapa con un avance de 6 meses aproximadamente. Sin embargo, también se pueden ver dos advertencias del modelo donde no hubo registro de fallo siendo FPs.

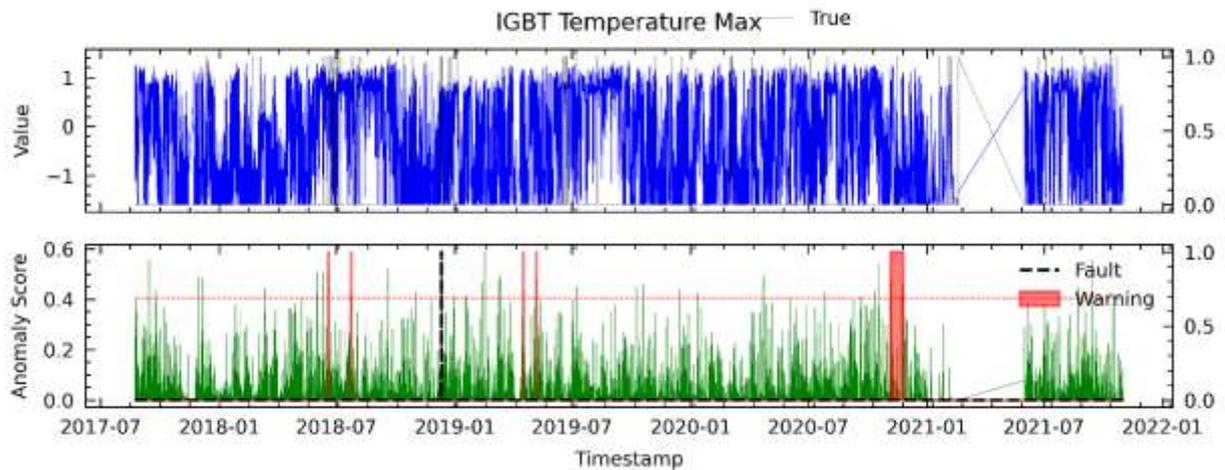


Figura 80. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

B. Anomaly Transformer utilizando el método RFE12 en WT6

En la **Figura 81** se presenta el resultado del análisis de la falla del módulo IGBT con las variables seleccionadas en RFE12, donde se observa que el modelo no advirtió la falla registrada del IGBT. Por otro lado, produjo un FP.

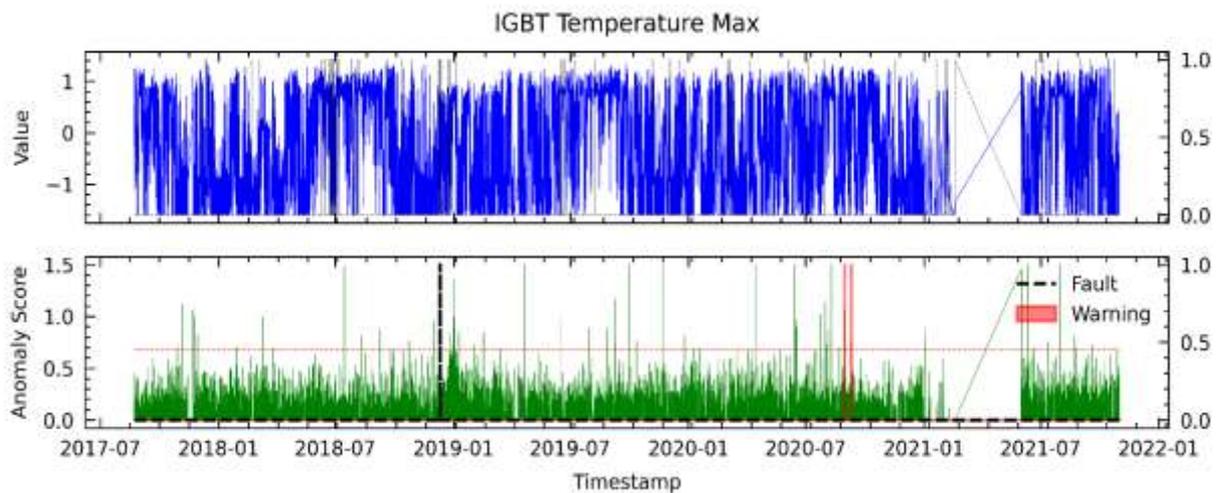


Figura 81. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

C. Anomaly Transformer utilizando el método RFE24 en WT6

En la **Figura 82** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que no se detectó la falla, pero en cambio se registra 2 FPs.

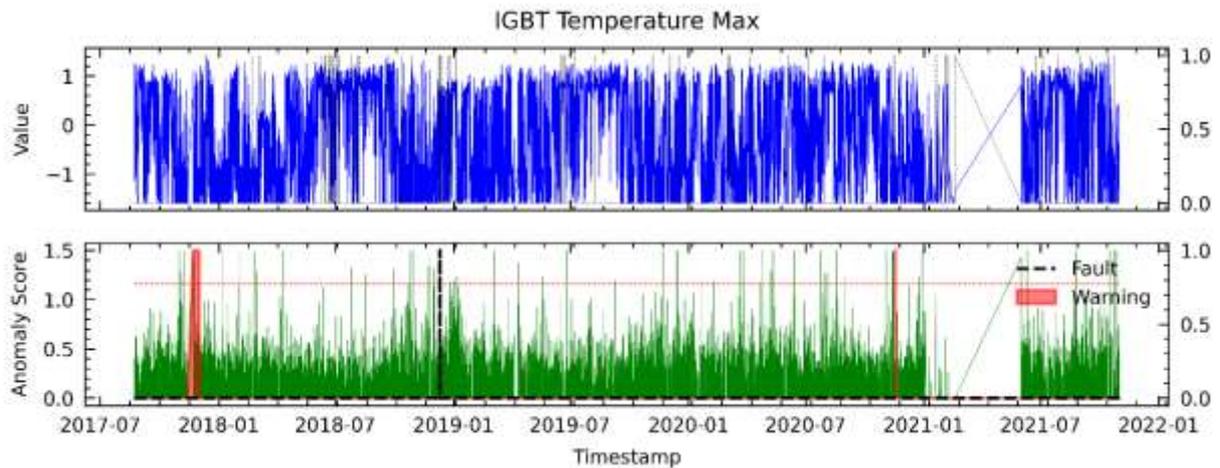


Figura 82. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

D. Anomaly Transformer utilizando el método C6 en WT6

En el método de selección de variables por correlación como entrada al modelo resulta en la **Figura 83**. Se puede observar que el modelo logró detectar la falla en el módulo IGBT con un avance aproximado de 1 semana de anticipación aproximadamente, pero también reportó tres FPs.

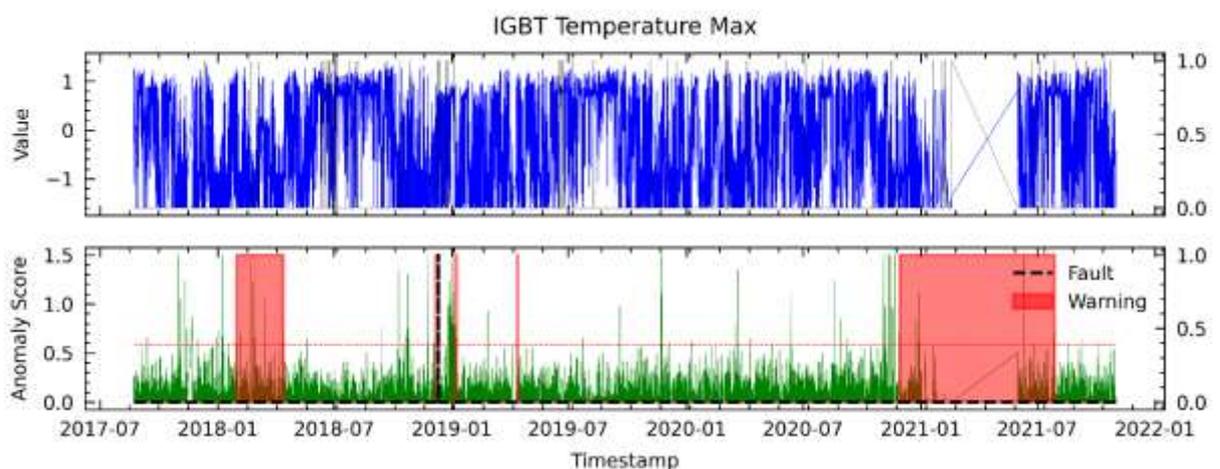


Figura 83. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

E. Anomaly Transformer utilizando el método PV18 en WT6

La **Figura 84** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió la falla del módulo IGBT con un avance de un mes aproximadamente y hubo un FP.

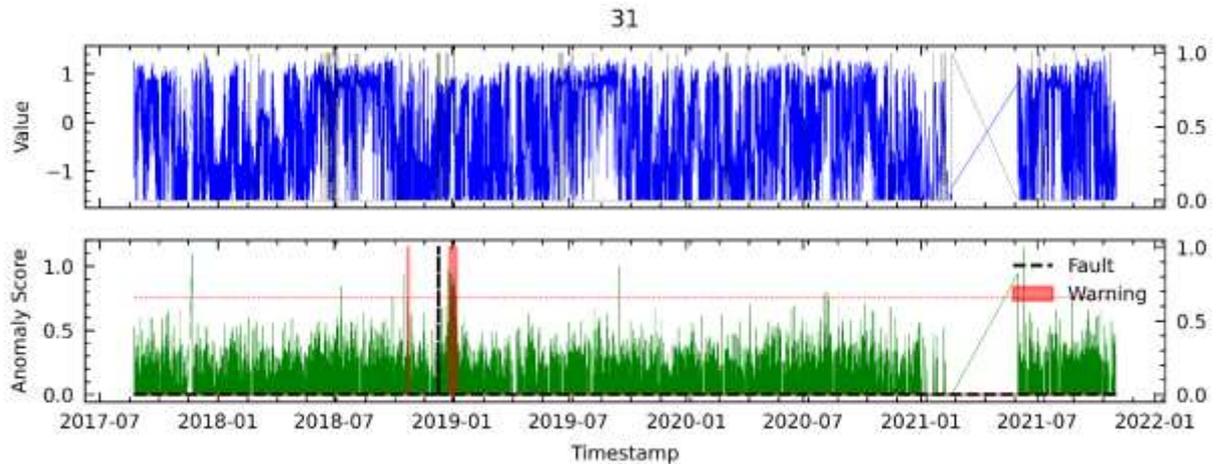


Figura 84. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

F. Anomaly Transformer utilizando el método AD23 en WT6

En la **Figura 85** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados son satisfactorios, ya que, es capaz de detectar la falla en el módulo IGBT con un avance aproximado de 4 meses de antelación.

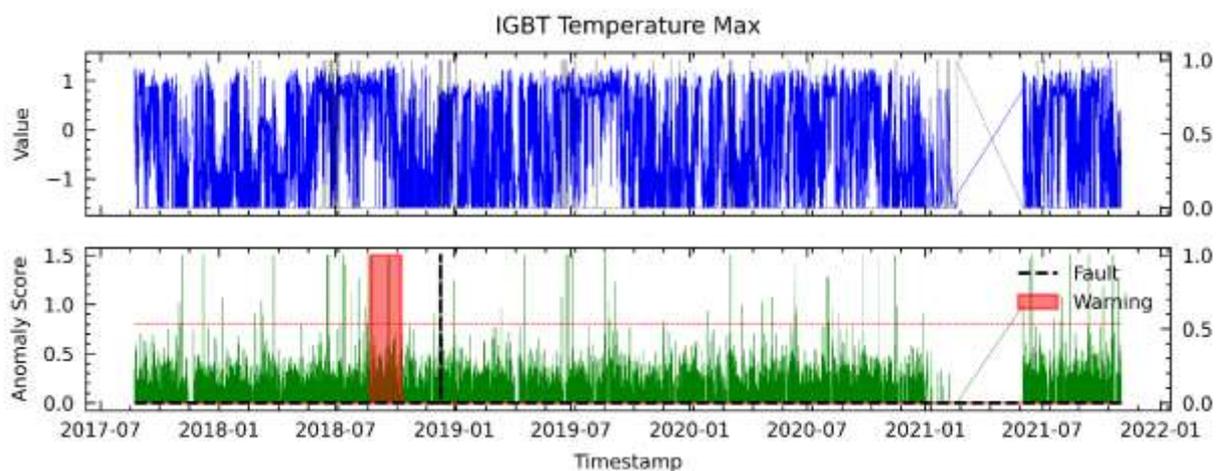


Figura 85. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

G. Anomaly Transformer utilizando el método LM18 en WT6

En la **Figura 86** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que se logró detectar la falla en el módulo IGBT con un avance aproximado de 2 meses de anticipación, además registró 1 FP.

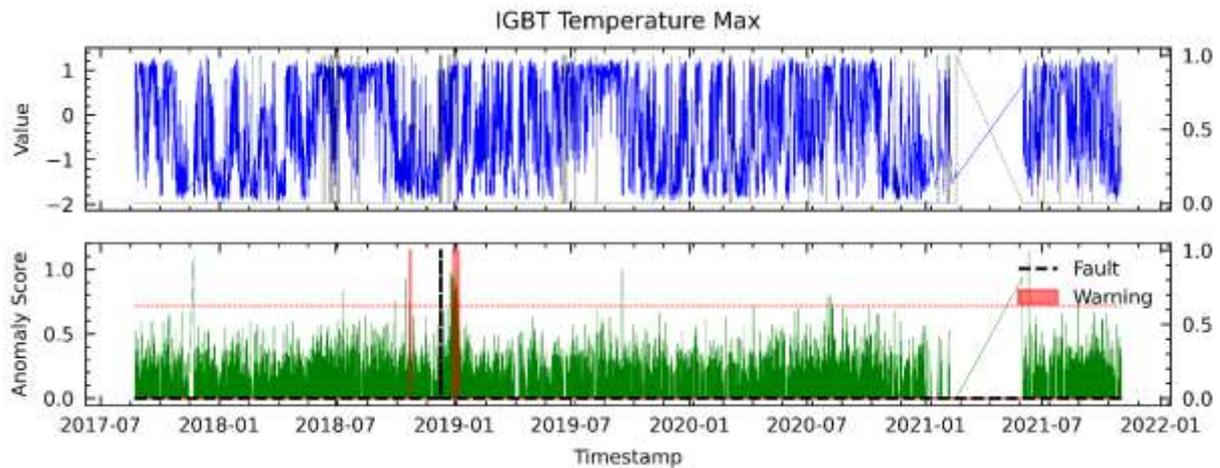


Figura 86. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT6.

VI. Detección de fallas en WT7 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT7

En la **Figura 87** se observa que el puntaje de anomalías supera el umbral de primera etapa con un avance promedio aproximado de 4 meses antes de la falla del componente IGBT, donde se detectaron las tres fallas registradas. Sin embargo, también se pueden observar cuatro advertencias del modelo donde no se encuentra registro de fallo del módulo IGBT.

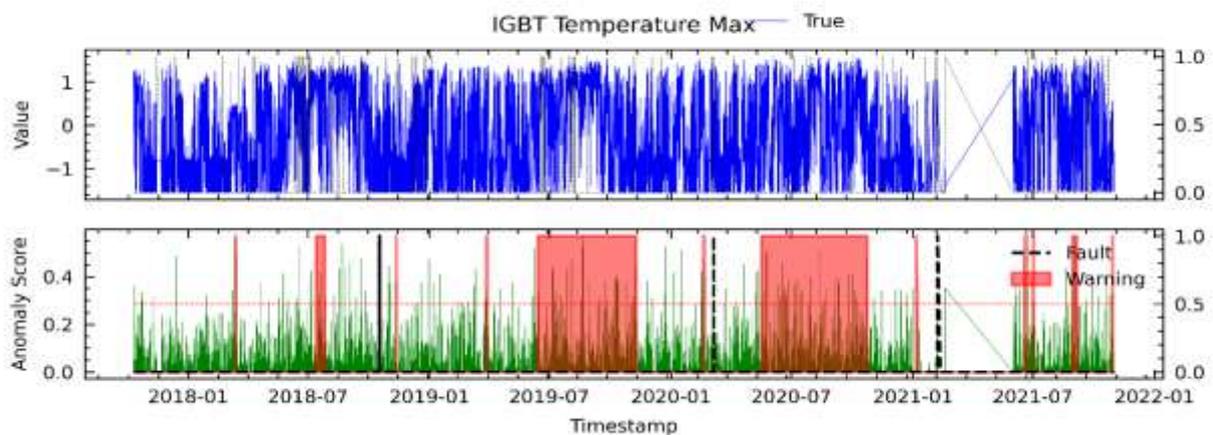


Figura 87. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

B. Anomaly Transformer utilizando el método RFE12 en WT7

En la **Figura 88** se presenta el resultado del análisis de la falla del módulo IGBT con el método RFE12, dando como resultado dos fallas advertidas con un avance aproximado de 6 meses de anticipación y dos FPs a finales de 2017 e inicios de 2018.

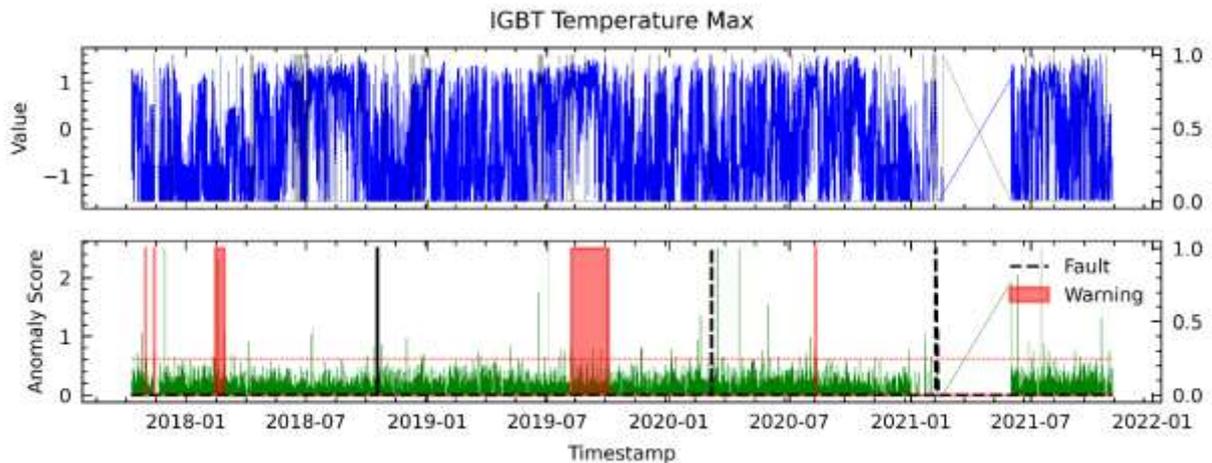


Figura 88. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

C. Anomaly Transformer utilizando el método RFE24 en WT7

En la **Figura 89** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RF24. Se observa que se detectó las 3 fallas, pero en cambio se registra 3 FPs.

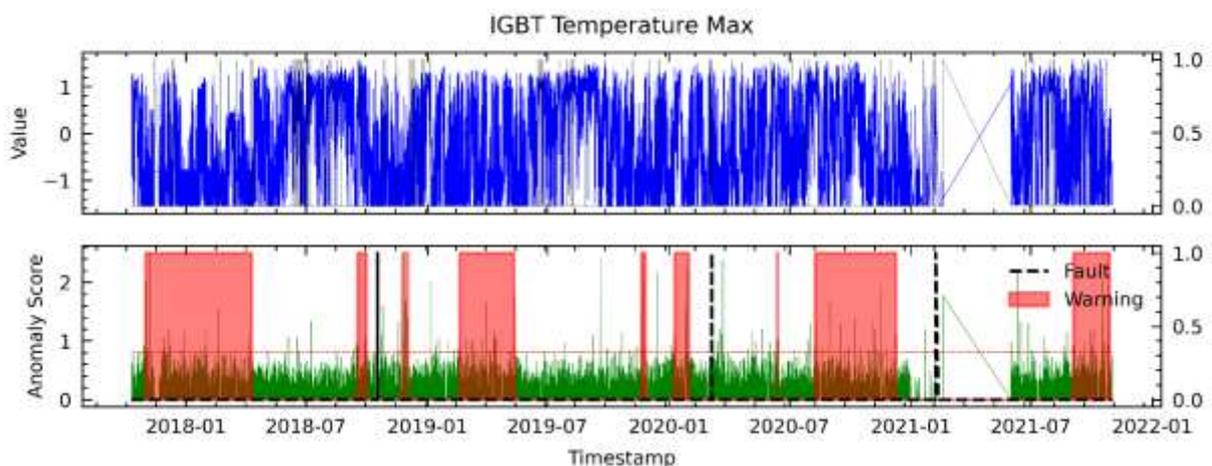


Figura 89. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

D. Anomaly Transformer utilizando el método C6 en WT7

En el método de selección de variables por correlación se eligieron variables como entrada al modelo, resultando en la **Figura 90**. Se puede observar que el modelo logró detectar las tres fallas del módulo IGBT con un avance aproximado de 6 meses de anticipación, pero también reportó 3 FPs.

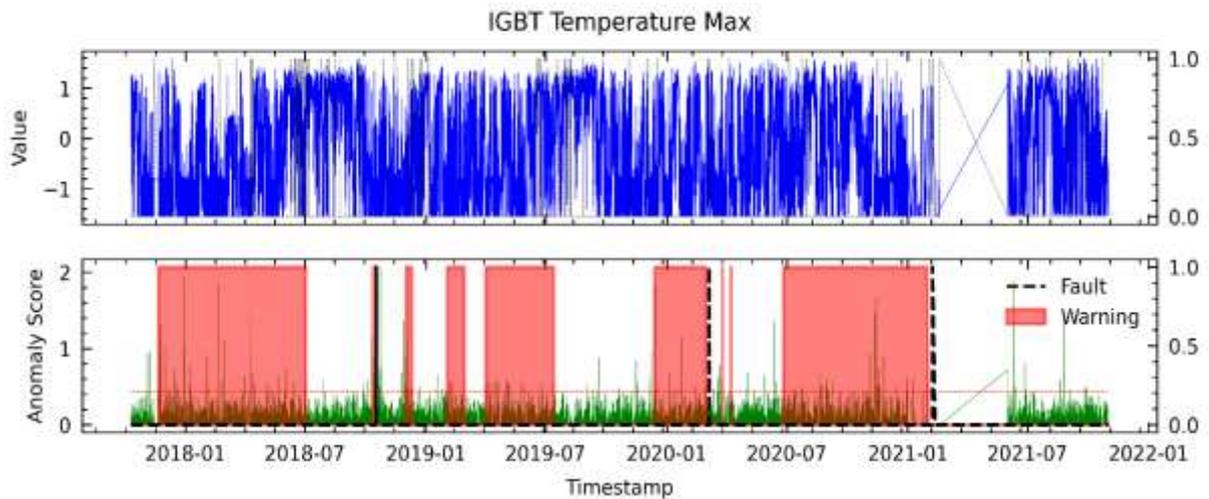


Figura 90. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

E. Anomaly Transformer utilizando el método PV18 en WT7

La **Figura 91** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió las tres fallas del módulo IGBT con un avance aproximado de 5 meses y registró un FP.

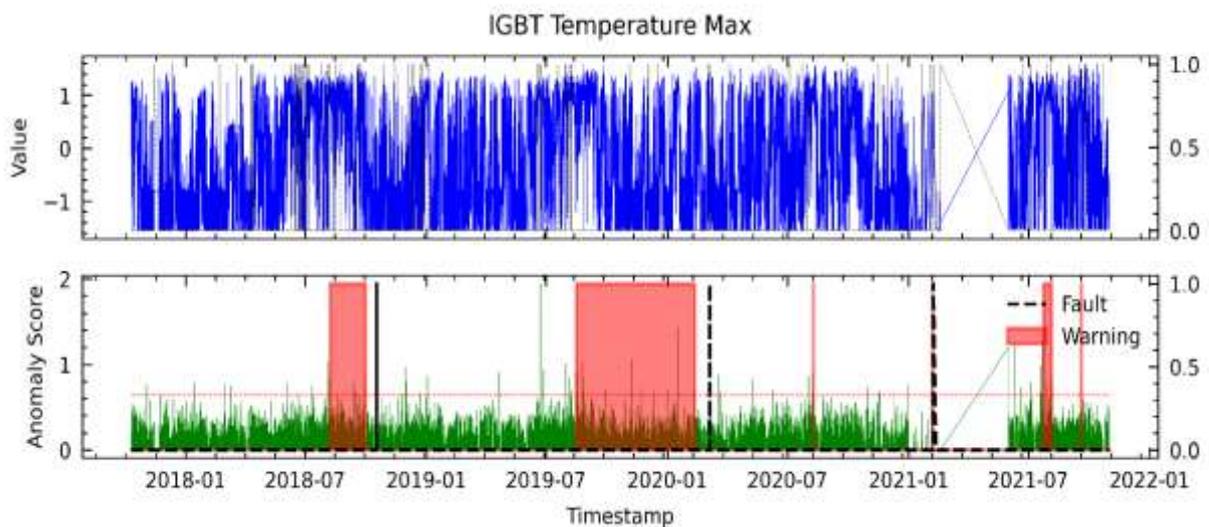


Figura 91. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

F. Anomaly Transformer utilizando el método AD23 en WT7

En la **Figura 92** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados no son satisfactorios, ya que, aunque es capaz de detectar las tres fallas en el módulo IGBT, también se han producido tres FPs.

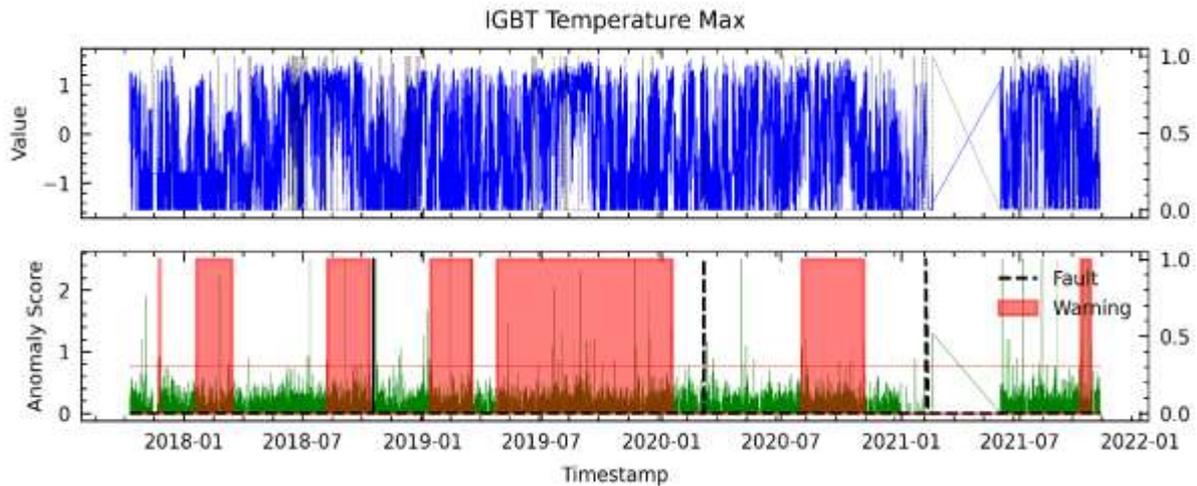


Figura 92. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

G. Anomaly Transformer utilizando el método LM18 en WT7

En la **Figura 93** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que se lograron detectar las tres fallas en el módulo IGBT con un avance aproximado de 5 meses de anticipación, pero también se presentó 1 FP.

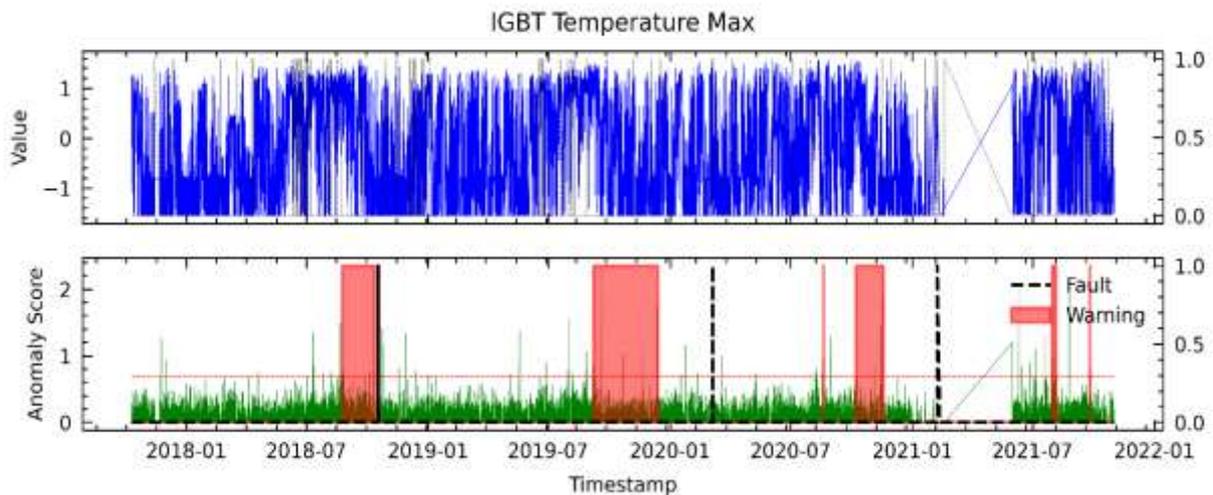


Figura 93. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT7.

VII. Detección de fallas en WT8 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT8

En la **Figura 94** se observa que el puntaje de anomalías supera el umbral de primera etapa con un avance aproximado de 6 meses antes de la falla del módulo IGBT.

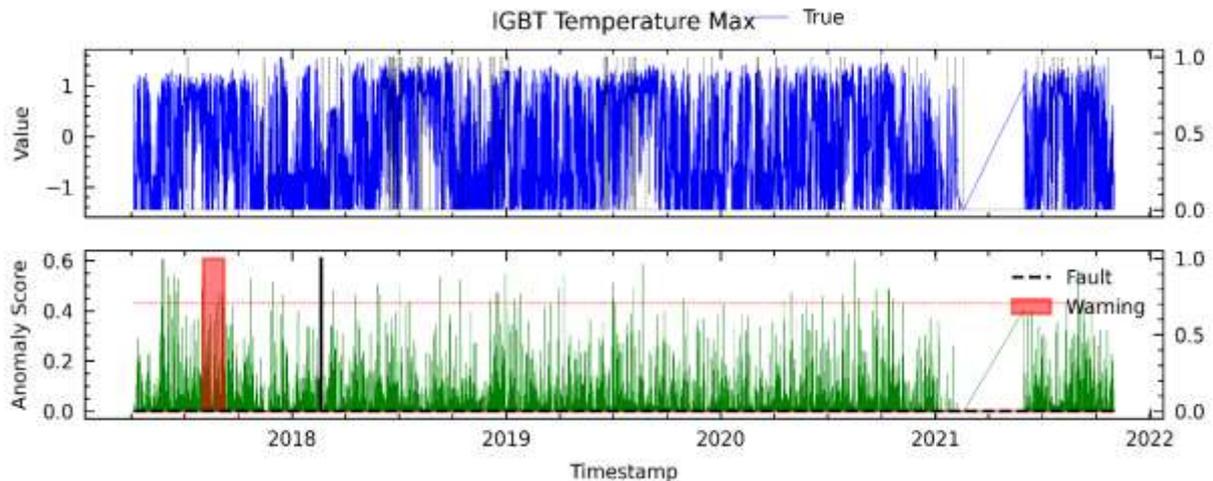


Figura 94. Resultados para RFE1 en el análisis de temperatura del módulo IGBT mediante el modelo AT para WT8.

B. Anomaly Transformer utilizando el método RFE12 en WT8

En la **Figura 95** se presenta el resultado del análisis de la falla del módulo IGBT con el método RFE12, dando como resultado que no se detectó la falla en el módulo IGBT y además registra 2 FPs.

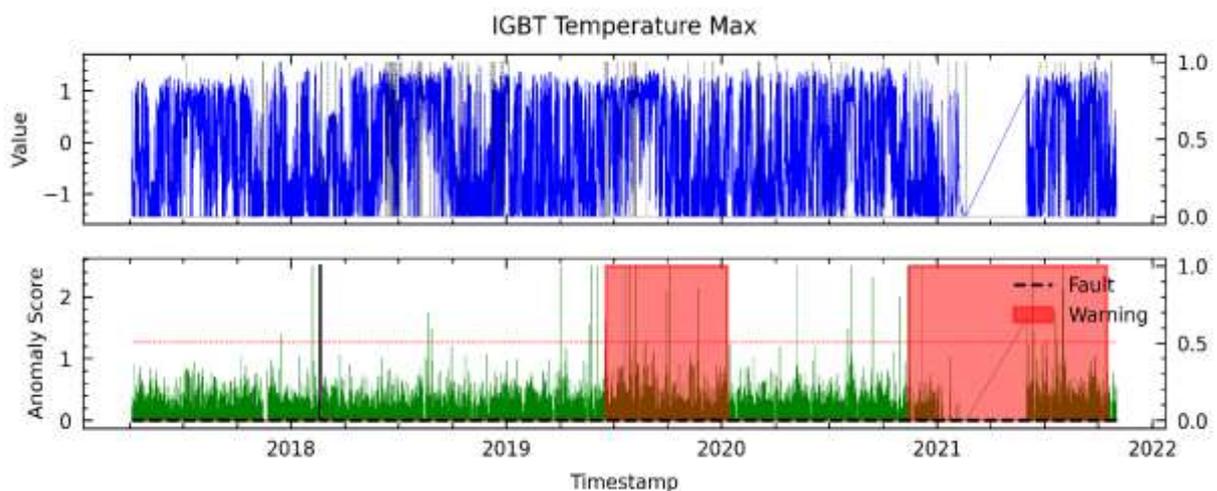


Figura 95. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.

C. Anomaly Transformer utilizando el método RFE24 en WT8

En la **Figura 96** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que no se detectó la falla, pero en cambio se registra 2 FPs similar al modelo RFE12.

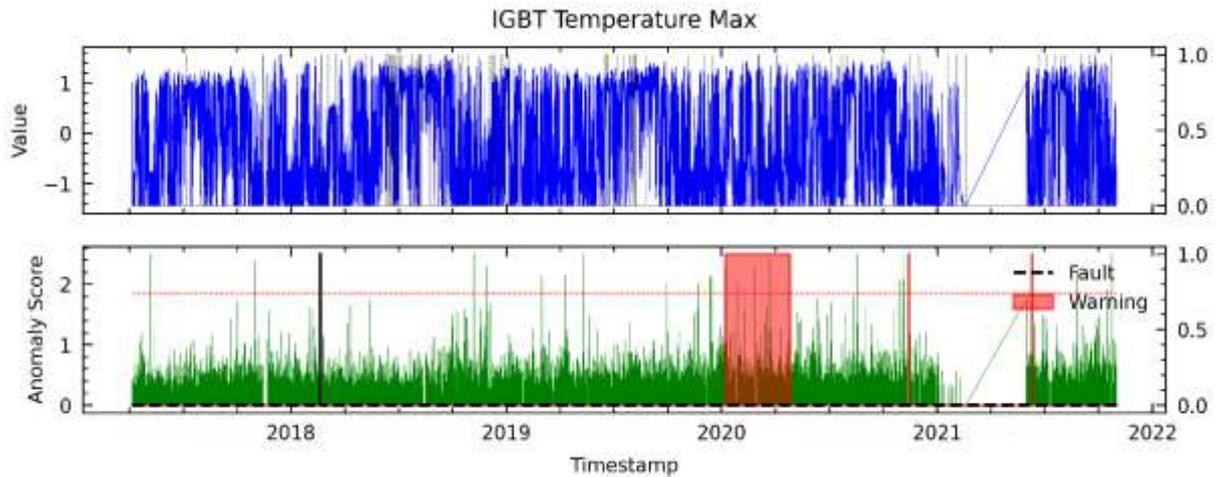


Figura 96. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.

D. Anomaly Transformer utilizando el método C6 en WT8

En el método de selección de características por correlación se eligieron variables como entrada al modelo, resultando en la **Figura 97**. Se puede observar que el modelo logró detectar la falla del módulo IGBT con un avance promedio de 2 meses aproximadamente, pero también reportó un FP que detectó una falla en el motor 1 de pitch registrado de O&M.

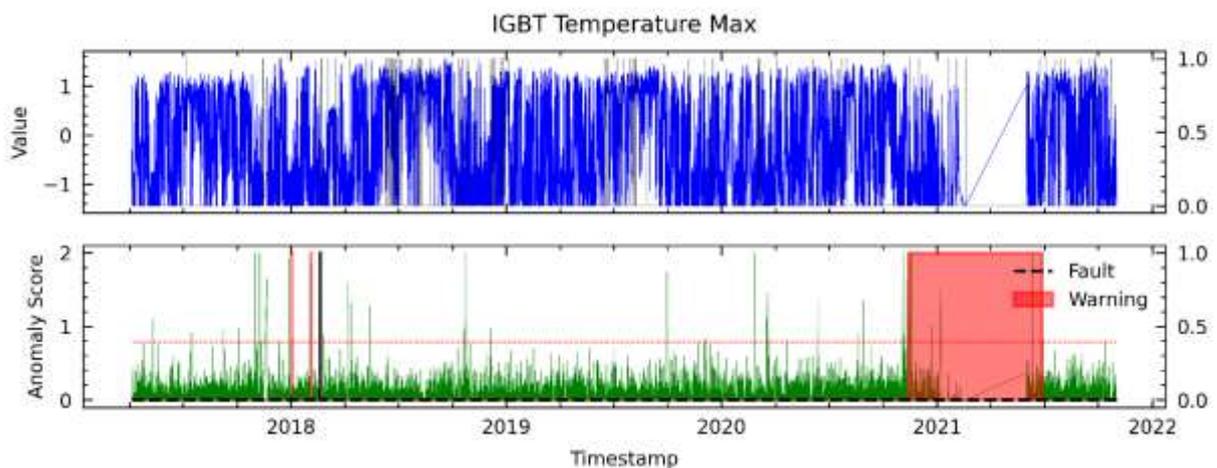


Figura 97. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.

E. Anomaly Transformer utilizando el método PV18 en WT8

La **Figura 98** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que el modelo no advirtió las tres fallas del módulo IGBT dando tres FP, donde uno corresponde a la falla del motor de pitch en el archivo de registro de O&M.

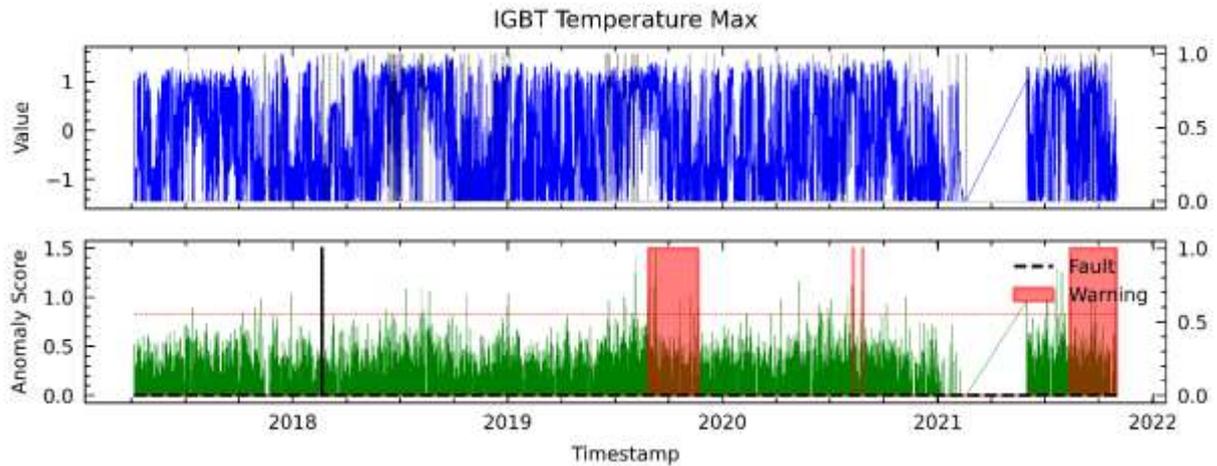


Figura 98. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.

F. Anomaly Transformer utilizando el método AD23 en WT8

En la **Figura 99** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados no son satisfactorios, ya que no es capaz de detectar las falla del IGBT. Sin embargo, se registra un FP.

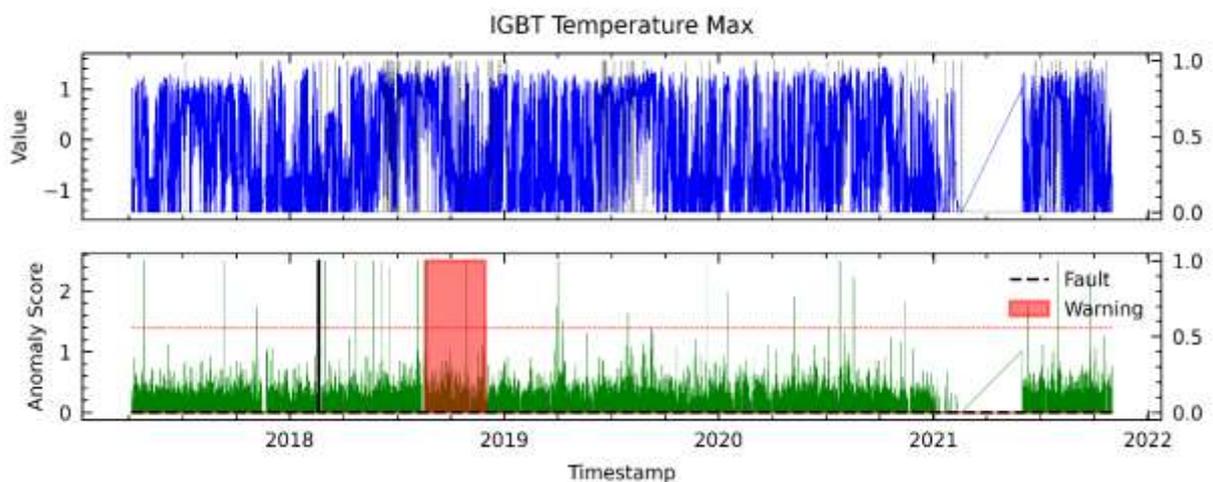


Figura 99. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.

G. Anomaly Transformer utilizando el método LM18 en WT8

En la **Figura 100** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que tampoco se logró detectar la falla en el módulo IGBT por lo que arrojó dos FPs.

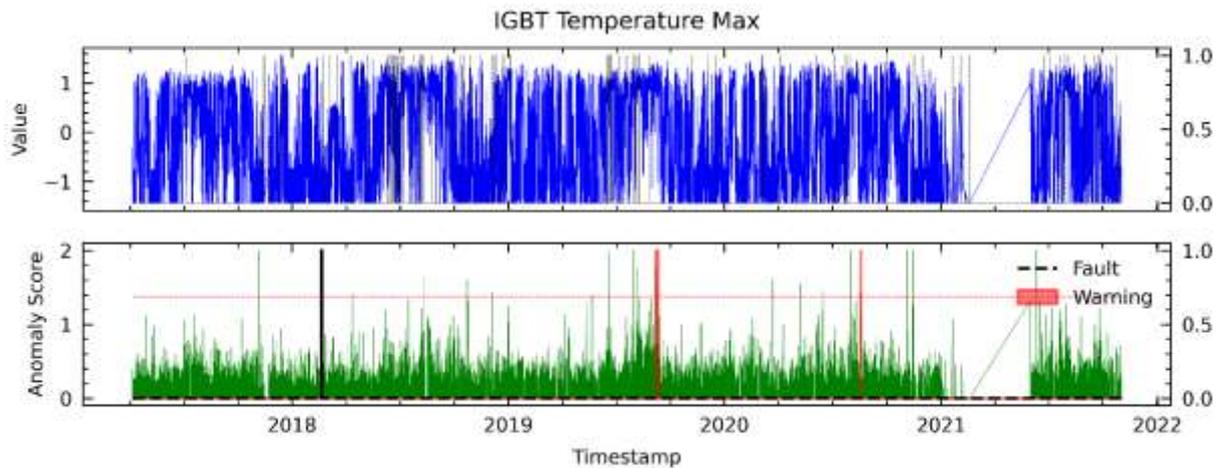


Figura 100. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT8.

VIII. Detección de fallas en WT9 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT9

En la **Figura 101** se observa que el puntaje de anomalías supera el umbral de primera etapa con un avance aproximado de 2 meses antes de la falla del módulo IGBT. Sin embargo, también se pueden ver cuatro advertencias del modelo donde no se encuentra registro de algún fallo.

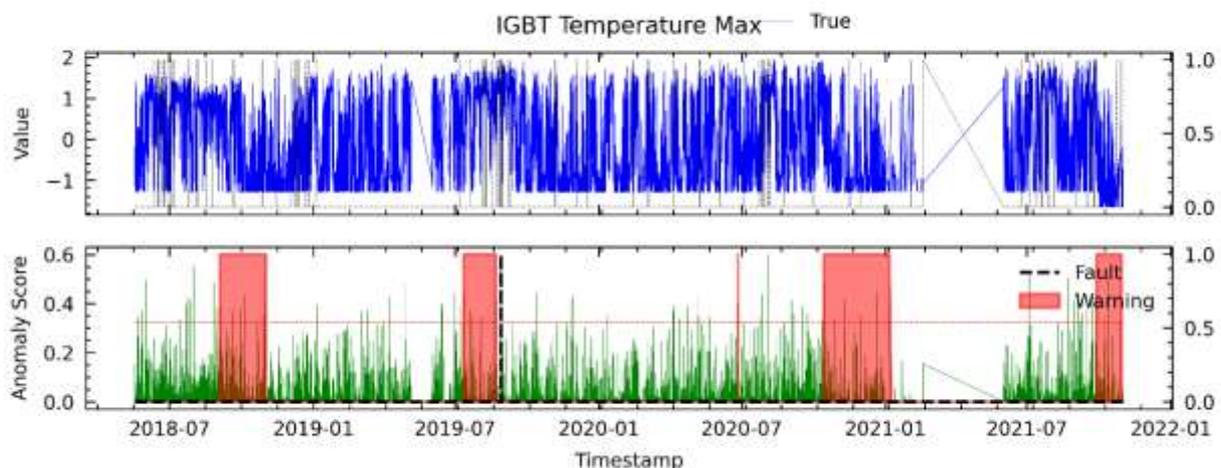


Figura 101. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.

B. Anomaly Transformer utilizando el método RFE12 en WT9

En la **Figura 102** se presenta el resultado del análisis de la falla del módulo IGBT con las variables de RFE12, dando como resultado la falla detectada con un avance aproximado de 2 semanas de anticipación y un FP a finales de 2021.

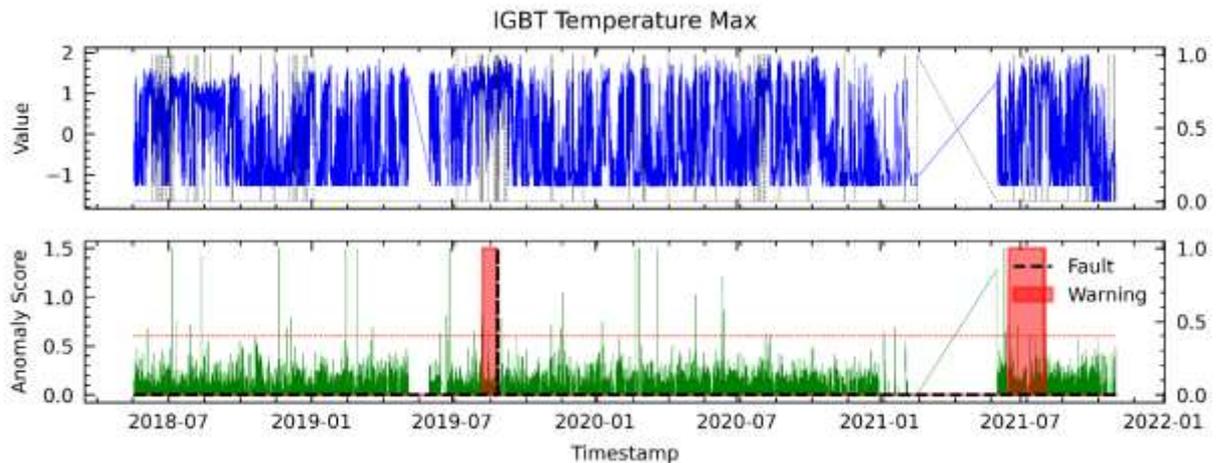


Figura 102. Resultados para RFE12 en el análisis de fallas del módulo mediante el modelo AT para WT9.

C. Anomaly Transformer utilizando el método RFE24 en WT9

En la **Figura 103** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que no se detectó la falla registrando 3 FPs.

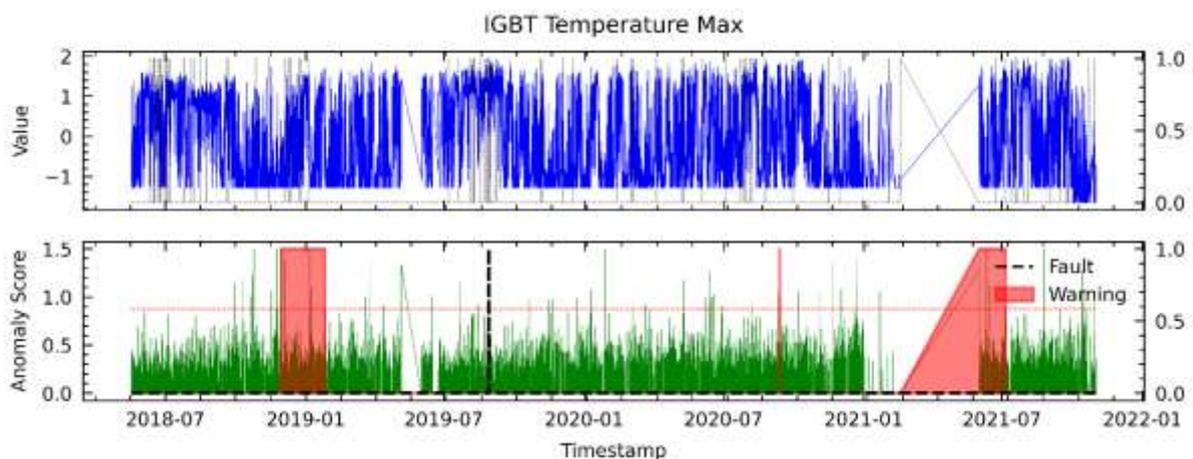


Figura 103. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.

D. Anomaly Transformer utilizando el método C6 en WT9

En el método de selección de características por correlación resulta la **Figura 104**. Se puede observar que el modelo logró detectar la falla del módulo IGBT con un avance aproximado de 7 meses de anticipación, pero también reportó 1 FP a inicios de 2021.

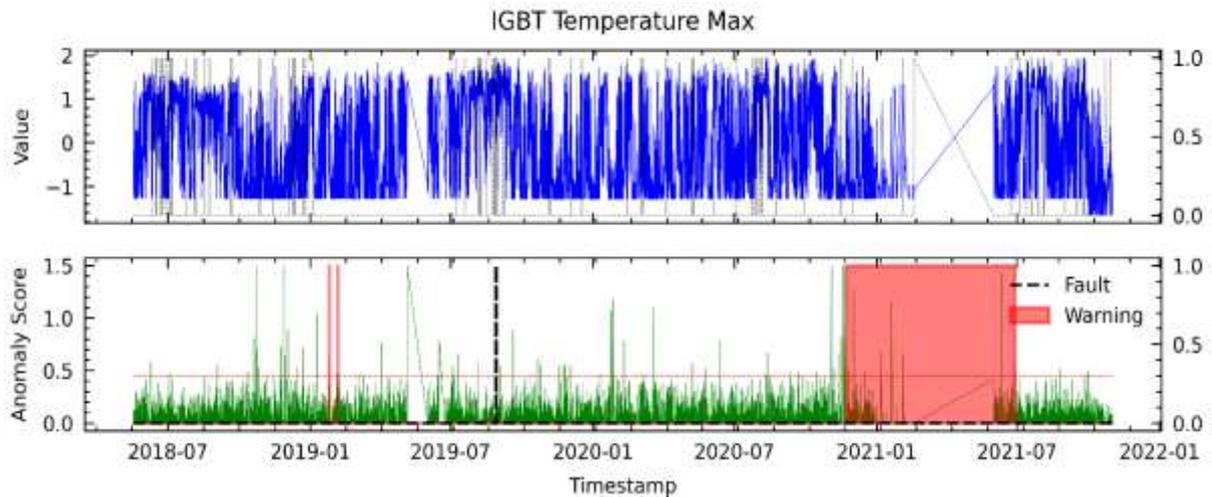


Figura 104. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.

E. Anomaly Transformer utilizando el método PV18 en WT9

La **Figura 105** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que el modelo advirtió la falla del módulo IGBT con un aproximado de 1 semana de anticipación y además hubo dos FPs.

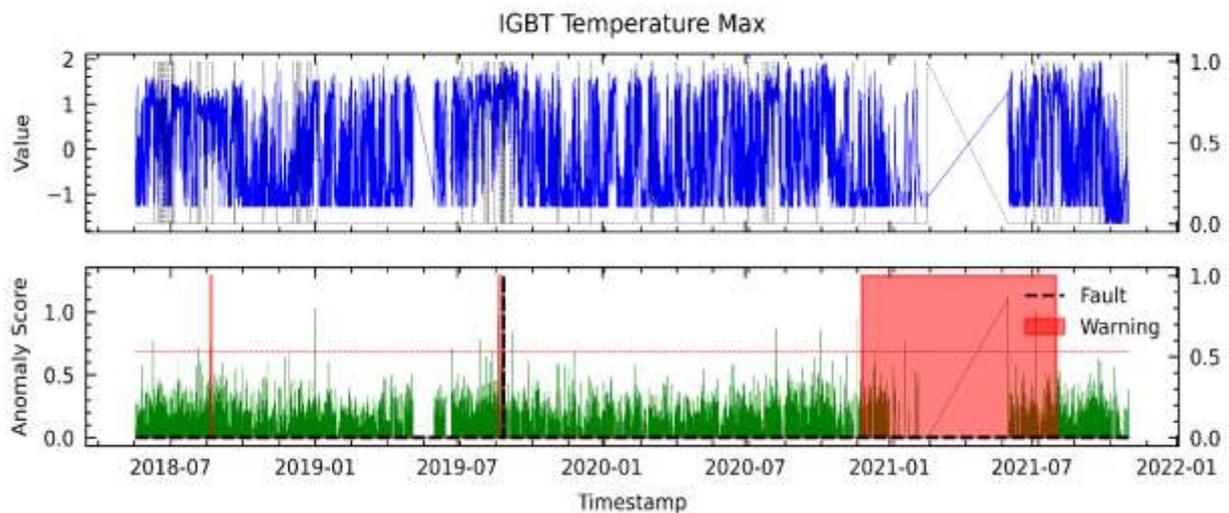


Figura 105. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.

F. Anomaly Transformer utilizando el método AD23 en WT9

En la **Figura 106** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. Los resultados son satisfactorios, ya que fue capaz de advertir con un avance aproximado de 6 meses antes de la falla en el módulo IGBT.

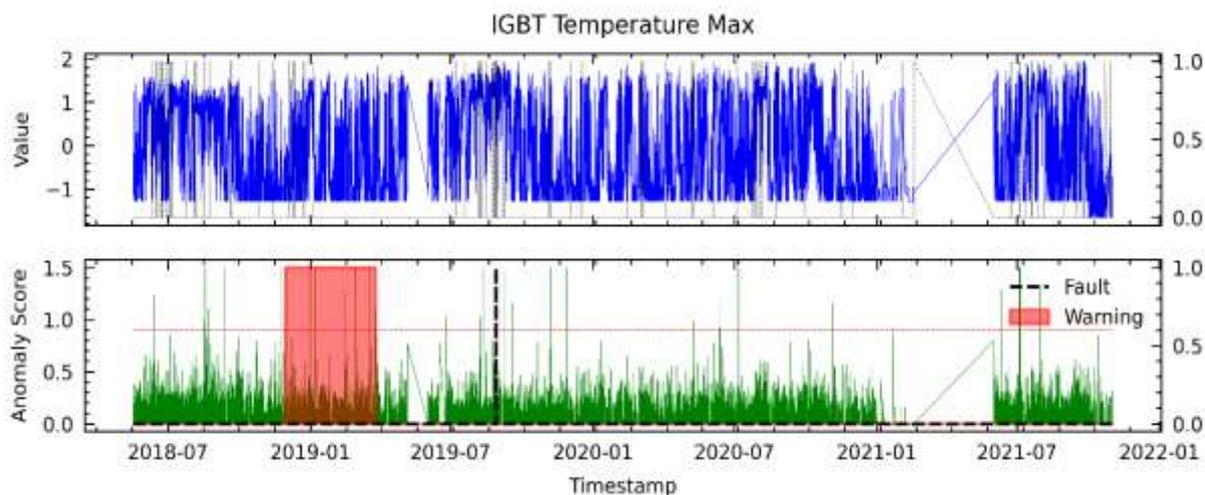


Figura 106. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.

G. Anomaly Transformer utilizando el método LM18 en WT9

En la **Figura 107** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que logró detectar la falla en el módulo IGBT con un aproximado de 1 mes de anticipación, pero también se presentó un FP.

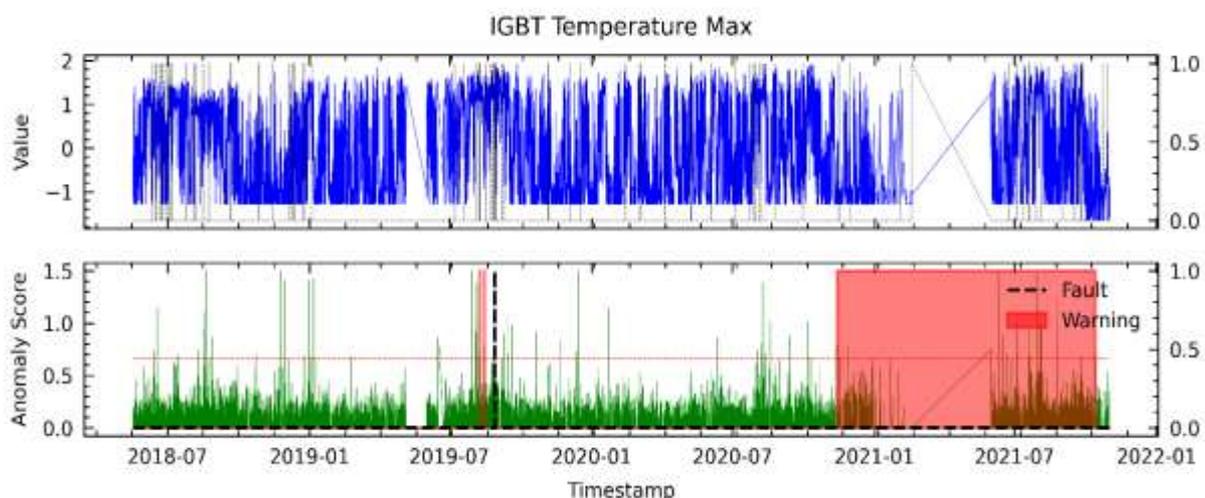


Figura 107. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT9.

IX. Detección de fallas en WT10 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT10

En la **Figura 108** se observa que el puntaje de anomalías supera el umbral de primera etapa incluso hasta 4 meses antes de la falla del componente IGBT, donde no se detectó una de las fallas registradas del módulo IGBT. Sin embargo, no es como se lo espera ya que el modelo detecta la segunda falla con un aproximado de 6 meses de anticipación.

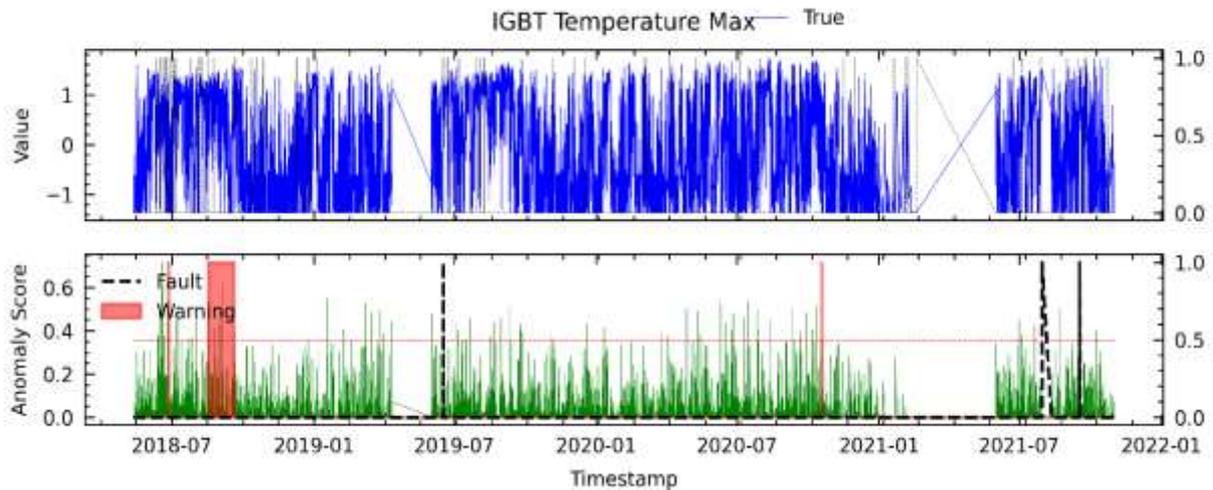


Figura 108. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.

B. Anomaly Transformer utilizando el método RFE12 en WT10

En la **Figura 109** se presenta el resultado del análisis de la falla del módulo IGBT con el método RFE12, dando como resultado una falla detectada con un avance aproximado de 7 meses de anticipación en la primera falla; además registra un FP y un FN.

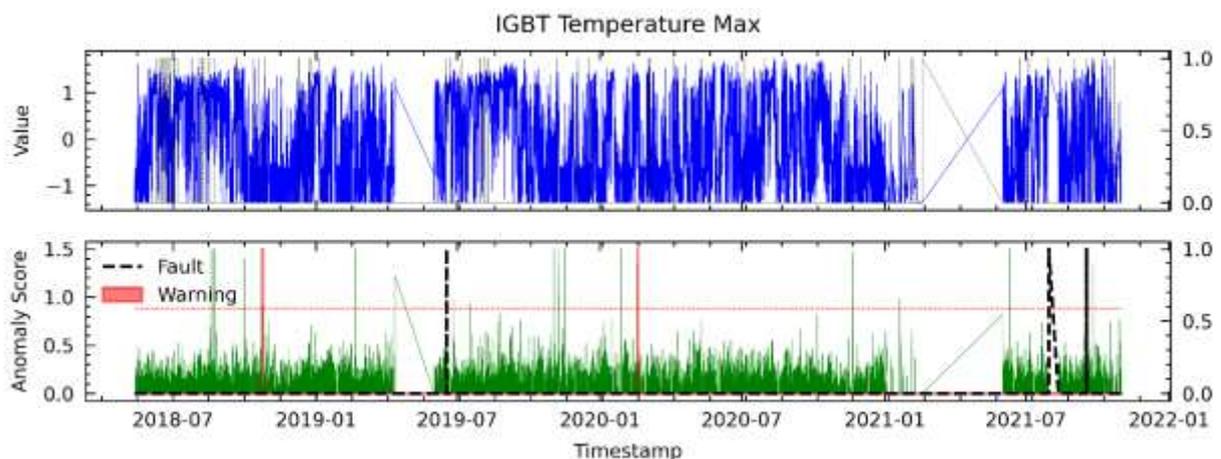


Figura 109. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.

C. Anomaly Transformer utilizando el método RFE24 en WT10

En la **Figura 110** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que se detectó las 2 fallas del módulo IGBT con un avance promedio aproximado de 4 meses antes de su ocurrencia.

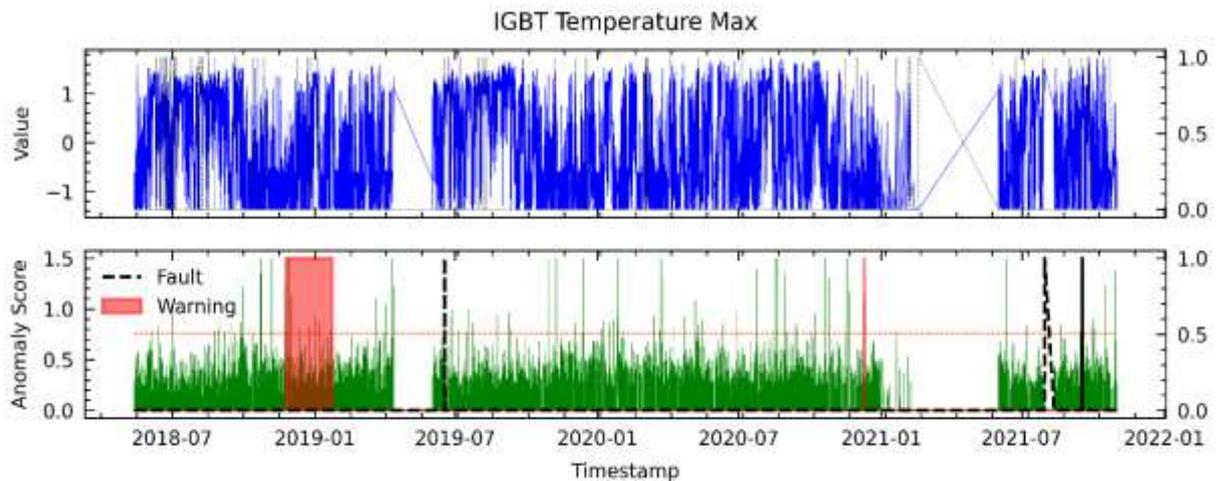


Figura 110. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.

D. Anomaly Transformer utilizando el método C6 en WT10

En el método de selección de características por correlación resulta en la **Figura 111**. Se puede observar que el modelo logró detectar las dos fallas del módulo IGBT con un avance promedio aproximado de 6 meses de anticipación.

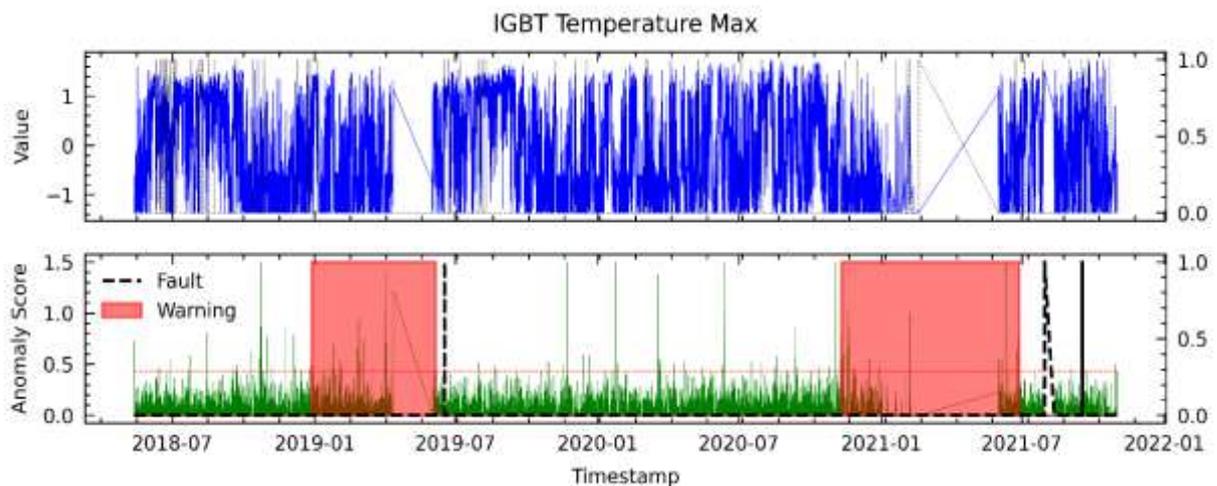


Figura 111. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.

E. Anomaly Transformer utilizando el método PV18 en WT10

La **Figura 112** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió las dos fallas del módulo IGBT en un aproximado de 8 meses y 1 semana respectivamente, y además registra un FP.

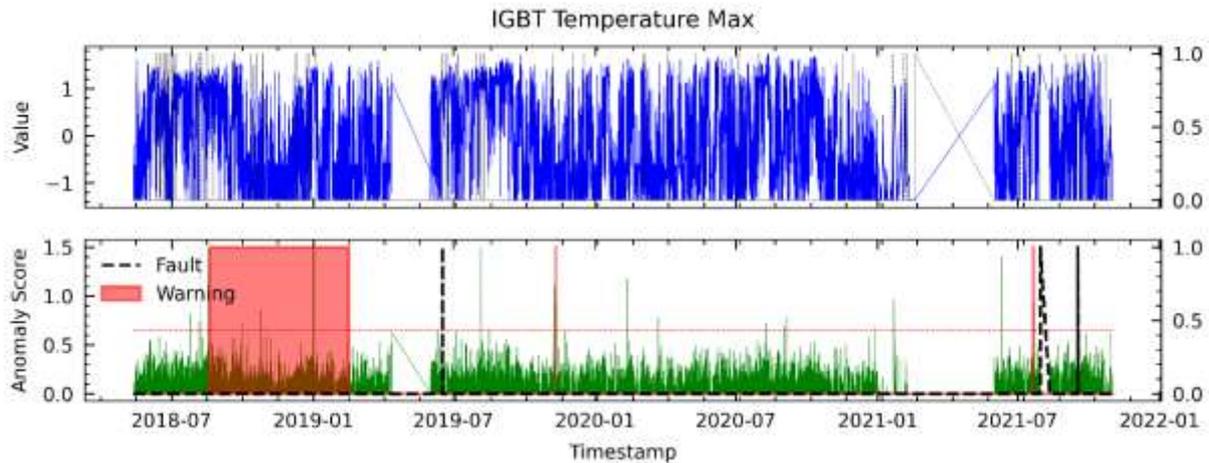


Figura 112. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.

F. Anomaly Transformer utilizando el método AD23 en WT10

En la **Figura 113** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23, donde se observa que el modelo es capaz de detectar las dos fallas en el módulo IGBT con un avance aproximado de 6 meses de anticipación, también se ha producido un FP.

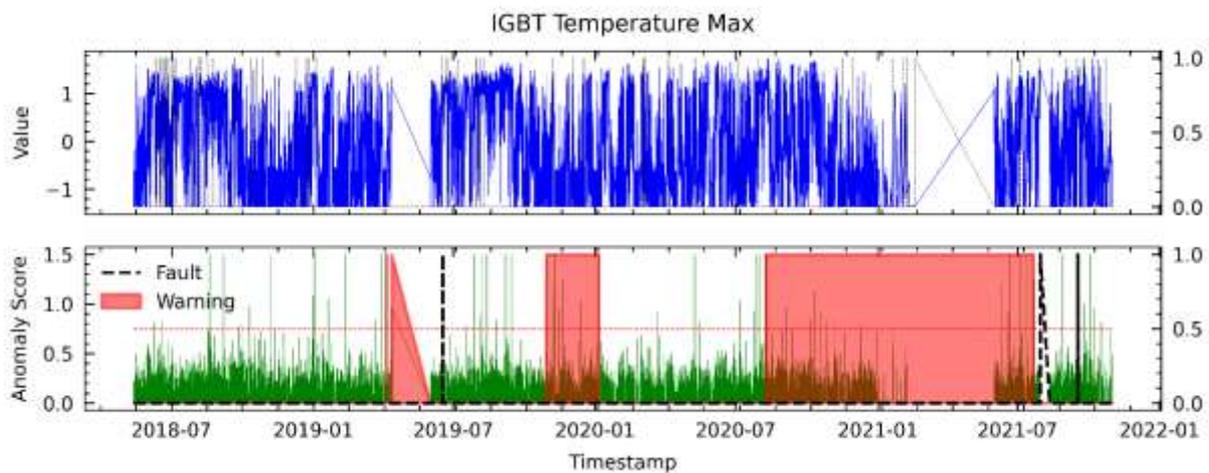


Figura 113. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.

G. Anomaly Transformer utilizando el método LM18 en WT10

En la **Figura 114** se muestran los resultados obtenidos con las variables seleccionadas por el método LM18. Se observa que se lograron detectar las 2 fallas en el módulo IGBT con un avance aproximado de 6 meses de anticipación, pero también se presentaron 1 FP.

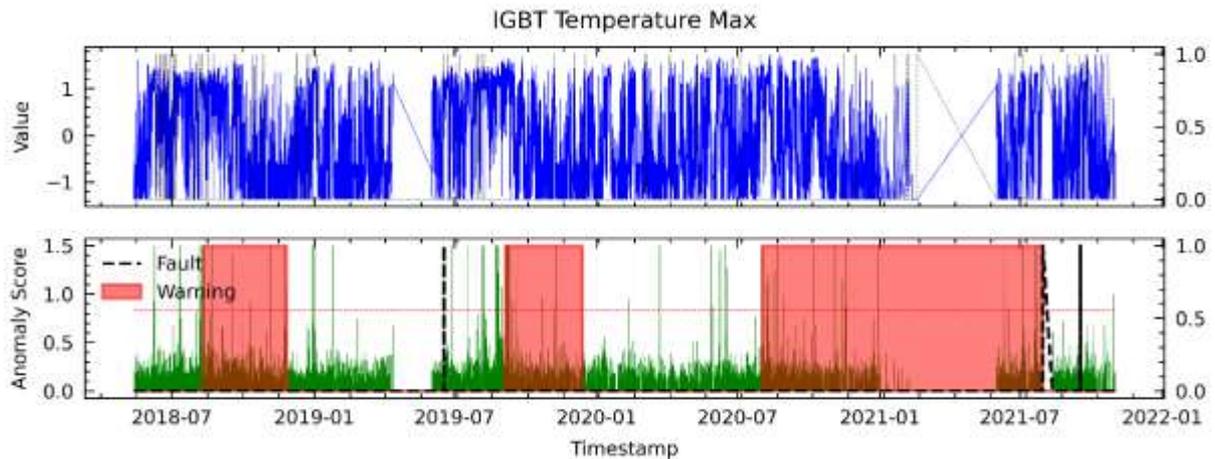


Figura 114. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT10.

X. Detección de fallas en WT11 mediante Anomaly Transformer (detalladamente).

A. Anomaly Transformer utilizando el método RFE1 en WT11

En la **Figura 115** se observa que el puntaje de anomalías supera el umbral de primera etapa con un avance aproximado de 3 meses antes de la falla del módulo IGBT. Sin embargo, también se pueden ver dos advertencias donde no se encuentra registro de algún fallo.

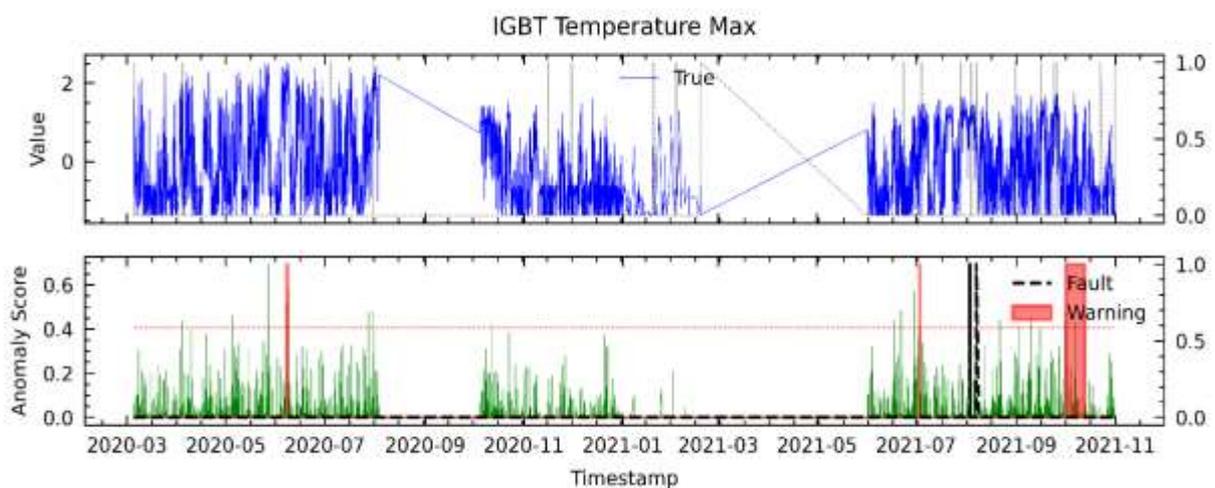


Figura 115. Resultados para RFE1 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

B. Anomaly Transformer utilizando el método RFE12 en WT11

En la **Figura 116** se presenta el resultado del análisis de la falla del módulo IGBT con el método RFE12, dando como resultado que se advirtió la falla con un avance aproximado de 3 meses de anticipación y también registró un FP a mediados de 2020.

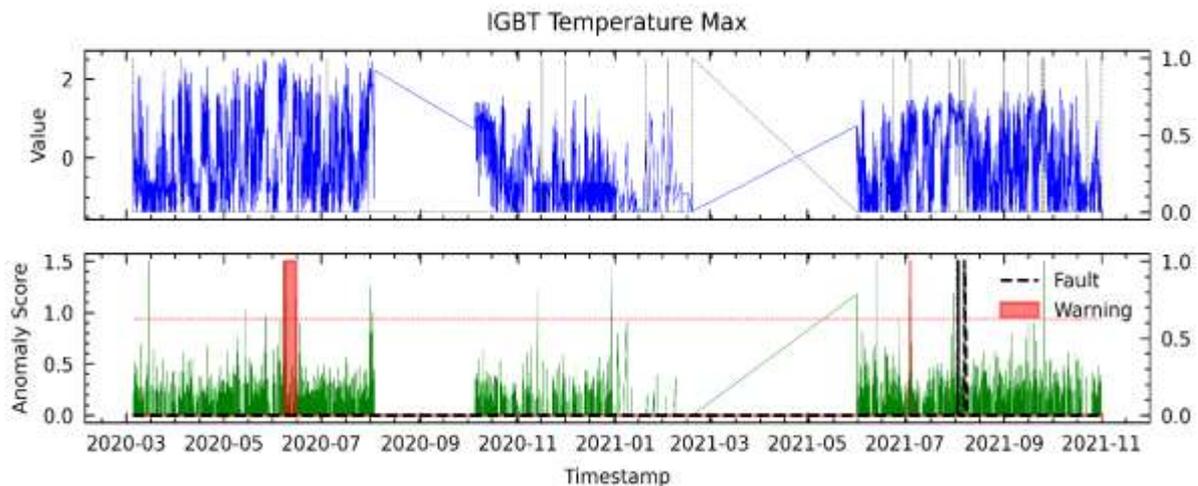


Figura 116. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

C. Anomaly Transformer utilizando el método RFE24 en WT11

En la **Figura 117** se muestra el resultado del modelo basado en las variables seleccionadas mediante el método RFE24. Se observa que el modelo detectó la falla con un avance aproximado de 8 meses antes, pero en cambio se registra un FP. También es posible notar que como existen huecos en la serie temporal es posible que no se detecte todos los patrones de falla.

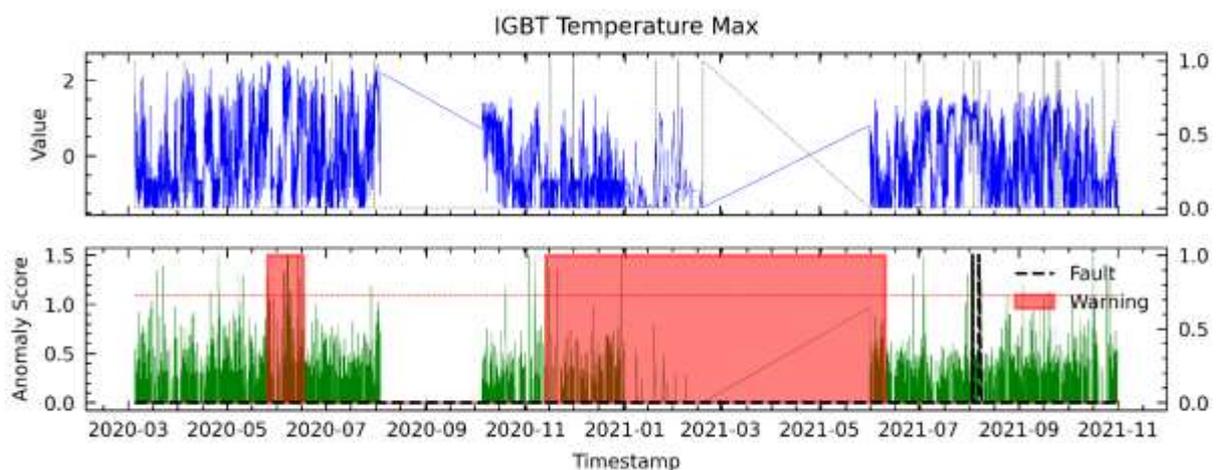


Figura 117. Resultados para RFE24 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

D. Anomaly Transformer utilizando el método C6 en WT11

En el método de selección de características por correlación se eligieron variables como entrada al modelo, resultando en la **Figura 118**. Se puede observar que el modelo logró detectar la falla con un avance aproximado de 5 meses antes de que la falla ocurriese en el módulo IGBT.

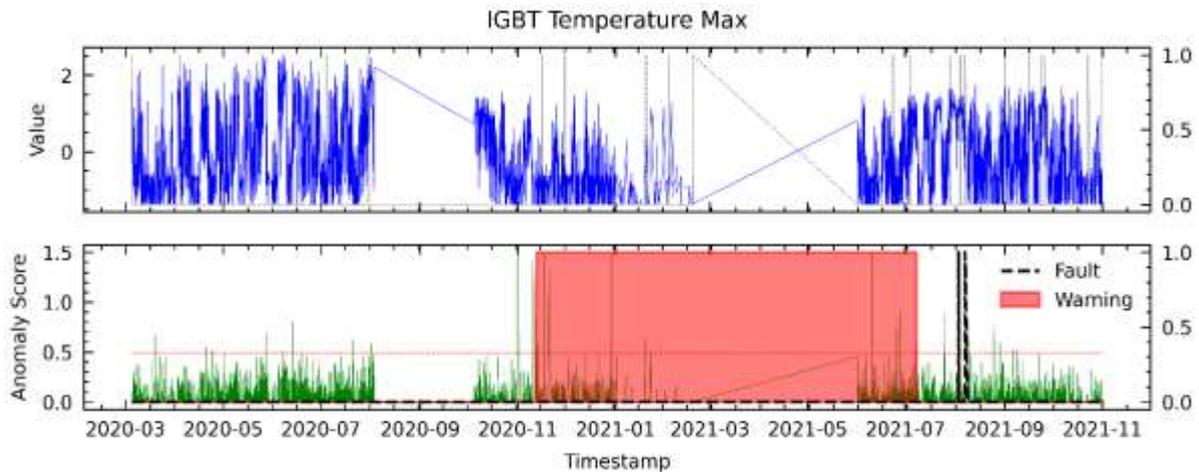


Figura 118. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

E. Anomaly Transformer utilizando el método PV18 en WT11

La **Figura 119** se describen los resultados obtenidos a partir de las variables seleccionadas utilizando el método PV18, donde se muestra que se advirtió la falla del módulo IGBT con un aproximado de 6 meses de anticipación, aunque no es posible argumentar debido a que existen huecos en la serie temporal.

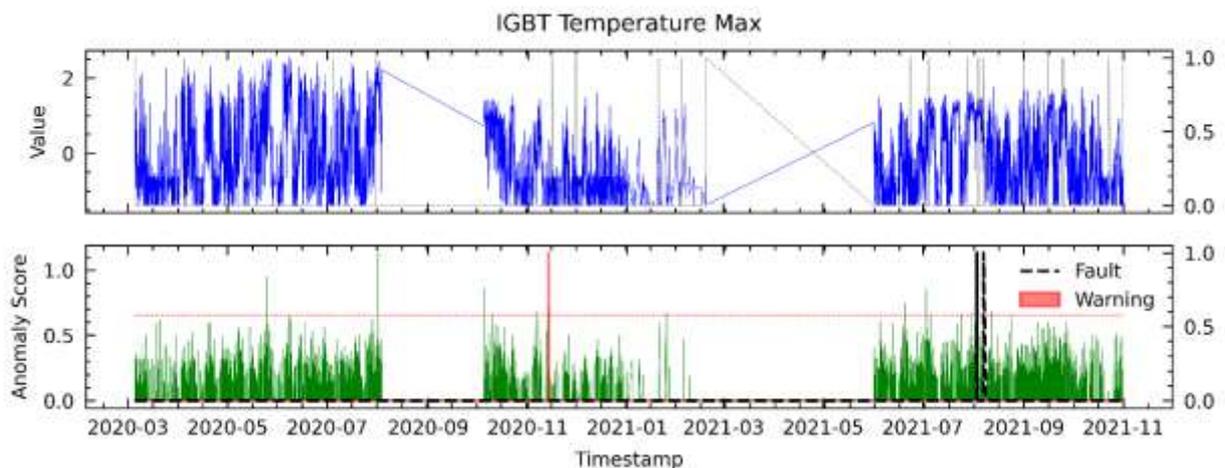


Figura 119. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

F. Anomaly Transformer utilizando el método AD23 en WT11

En la **Figura 120** se presentan los resultados obtenidos con las variables seleccionadas por el método AD23. El modelo es capaz de detectar la falla con un mes de anticipación, pero también registra un FP.

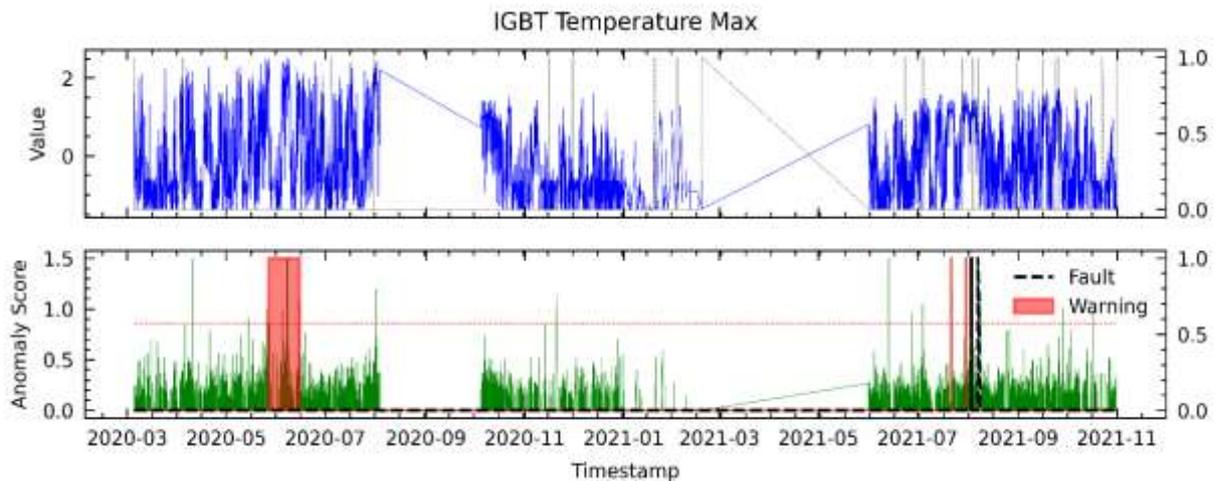


Figura 120. Resultados para AD23 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

G. Anomaly Transformer utilizando el método LM18 en WT11

En la **Figura 121** se muestran los resultados obtenidos con las variables seleccionadas por LM18. Se observa que se logró detectar la falla en el módulo IGBT con un aproximado de 1 semana de anticipación, pero también se presentaron dos FPs.

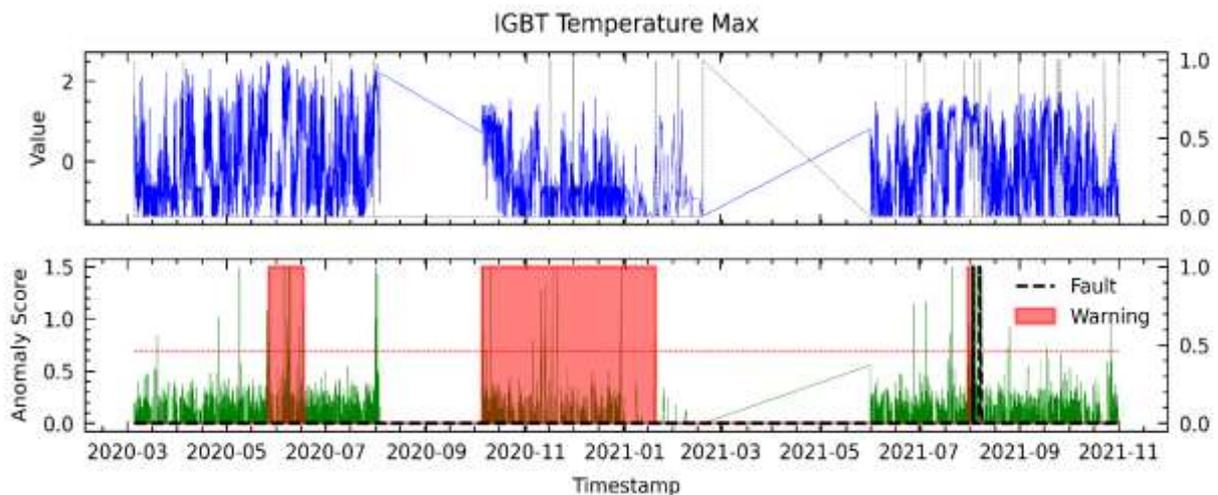


Figura 121. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo AT para WT11.

Anexo 4. Resultados para toda la selección de características desde la WT1 hasta la WT11 para TranAD

XI. Detección de fallas en WT1 mediante TranAD (detalladamente).

A. TranAD utilizando el método C6 en WT1

En la **Figura 122** se analiza variable a variable la evaluación de fallas de primera etapa. Las variables en análisis se graficaron con las variables del método C6.

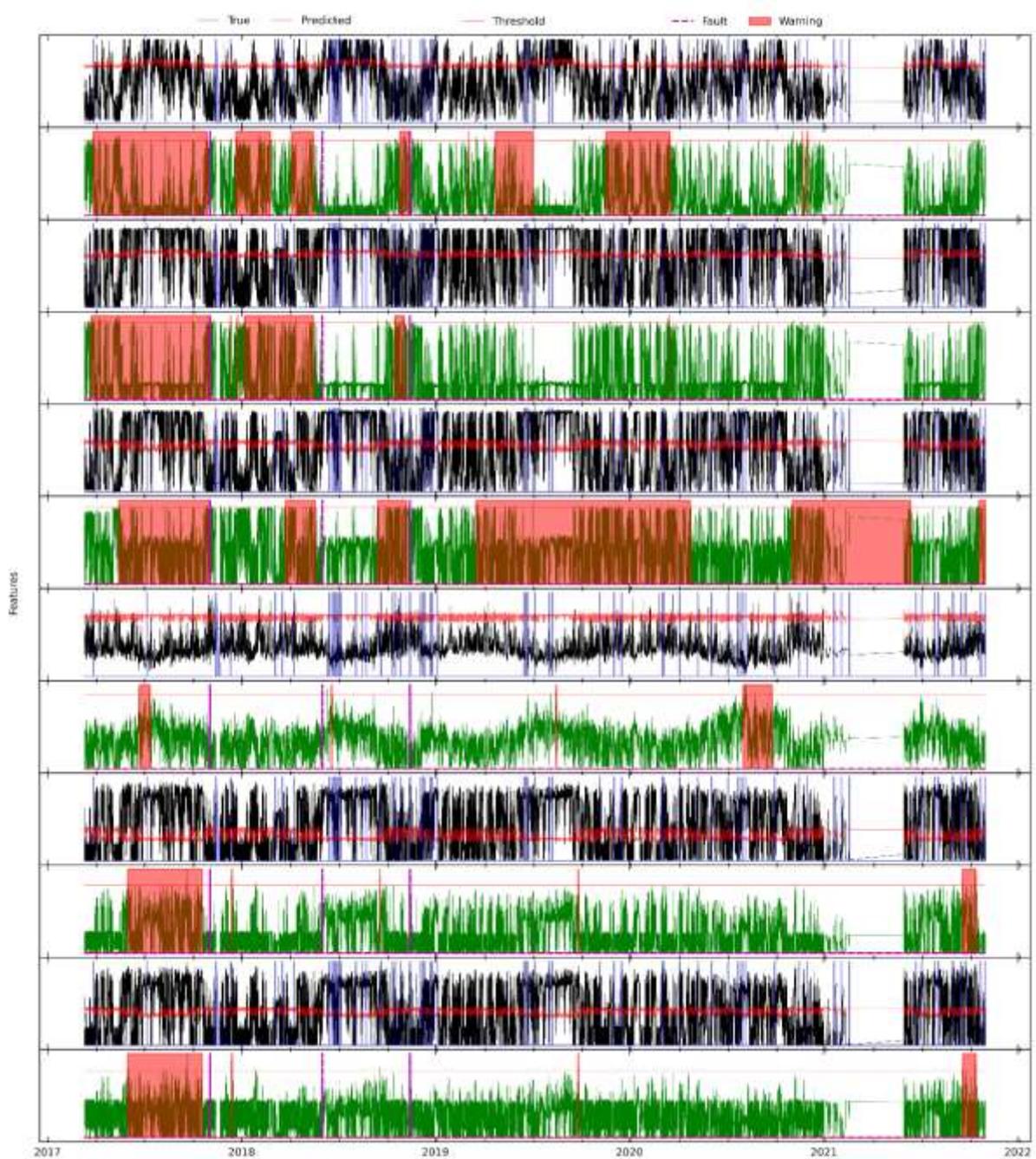


Figura 122. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT1.

En la **Figura 123** se observa el resultado de aplicar la evaluación de segunda etapa en la que se destaca que el modelo advirtió las tres fallas en el módulo IGBT con un avance promedio aproximado de 6, 2 y 1 meses respectivamente para cada falla.

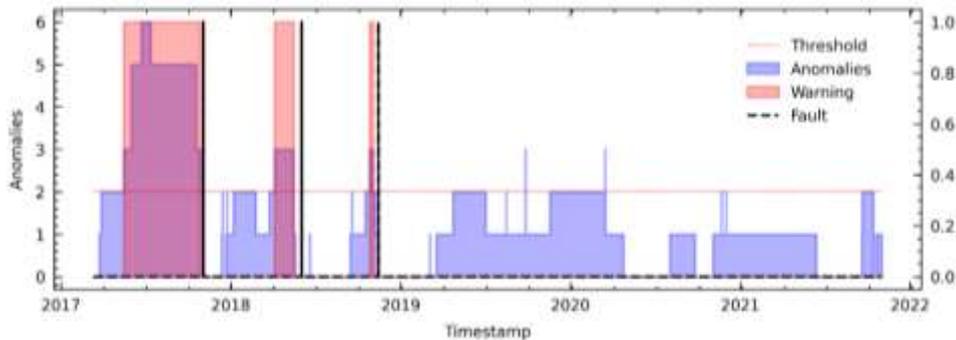


Figura 123. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.

B. TranAD utilizando el método PV18 en WT1

En la **Figura 124** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

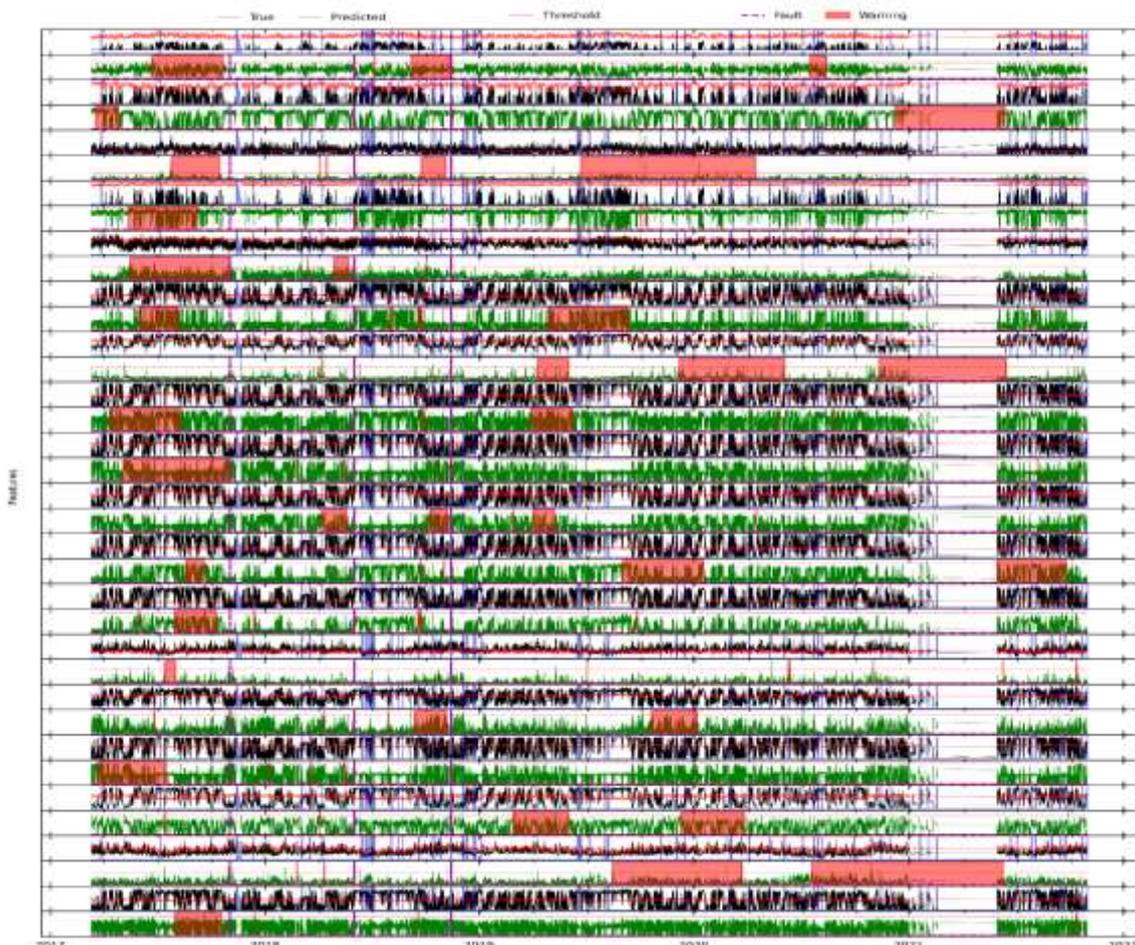


Figura 124. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT1.

En la **Figura 125** se muestra que el modelo solo logró advertir la primera falla en el módulo IGBT, por lo que en este caso se observó dos FNs.

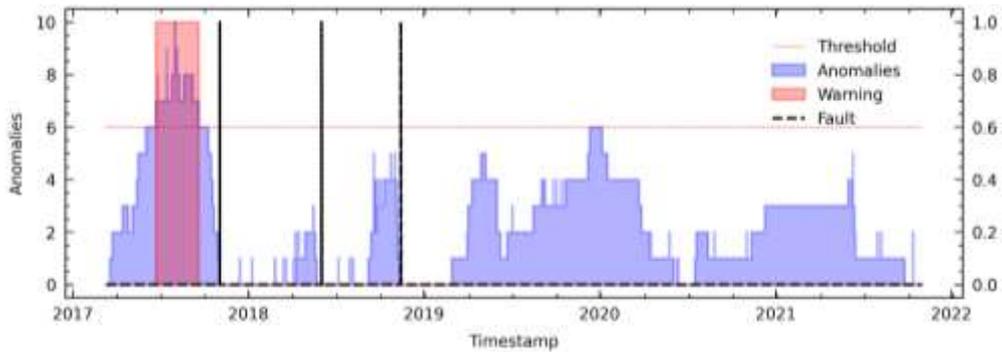


Figura 125. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.

C. TranAD utilizando el método LM18 en WT1

Para el análisis de LM18, en primer lugar, se debe analizar la **Figura 126**, donde se evalúa variable por variable.

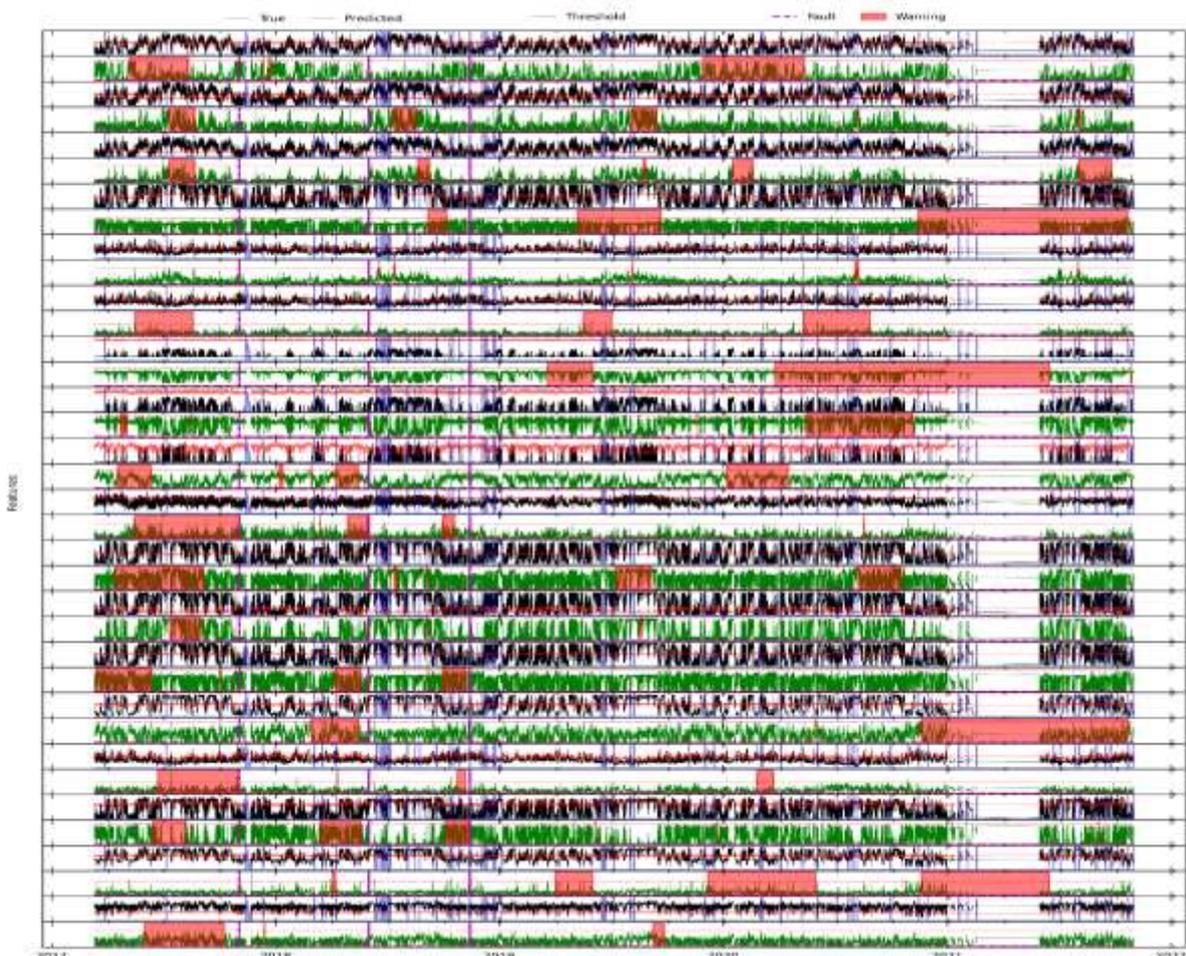


Figura 126. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT1.

En los resultados para LM18 se obtuvo la **Figura 127**, donde se observa que el modelo advirtió la primera falla aproximadamente 6 meses antes de su ocurrencia, así mismo la segunda falla con un avance de 2 meses aproximadamente. También se obtuvo un FN.

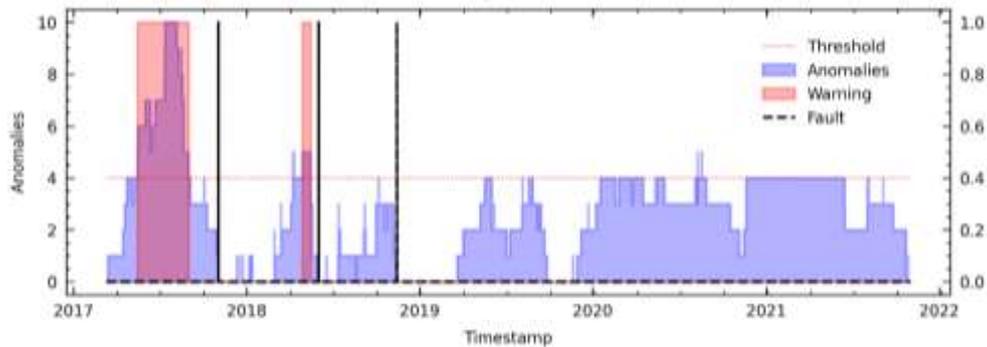


Figura 127. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT1.

XII. Detección de fallas en WT2 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT2

En la **Figura 128** se analiza variable a variable la evaluación de fallas.

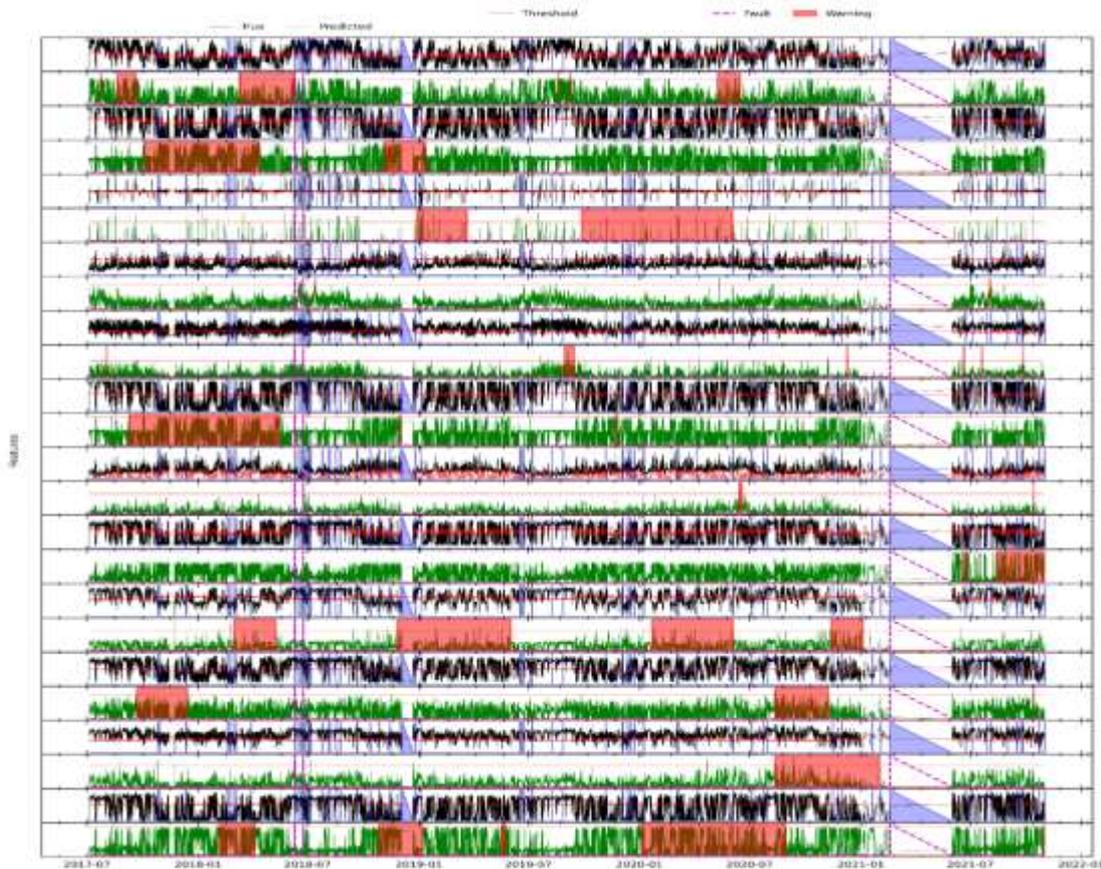


Figura 128. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.

En la **Figura 129** se destaca que el modelo advirtió las dos fallas con anticipación aproximada de 3 y 7 meses respectivamente. Además, se registra un FP a inicios de 2019.

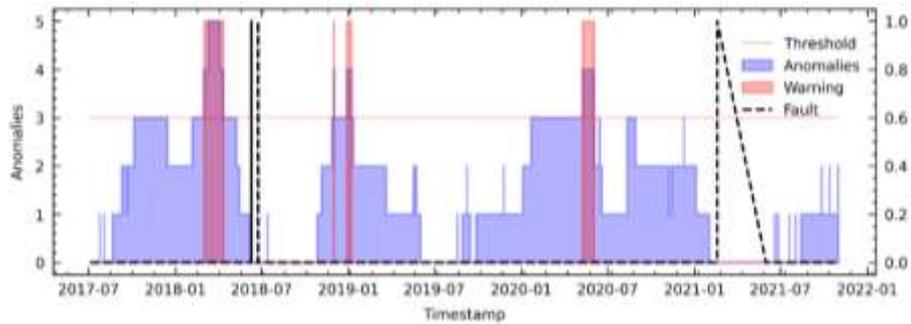


Figura 129. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.

B. TranAD utilizando el método C6 en WT2

En la **Figura 130** se analiza variable a variable la evaluación de advertencias.

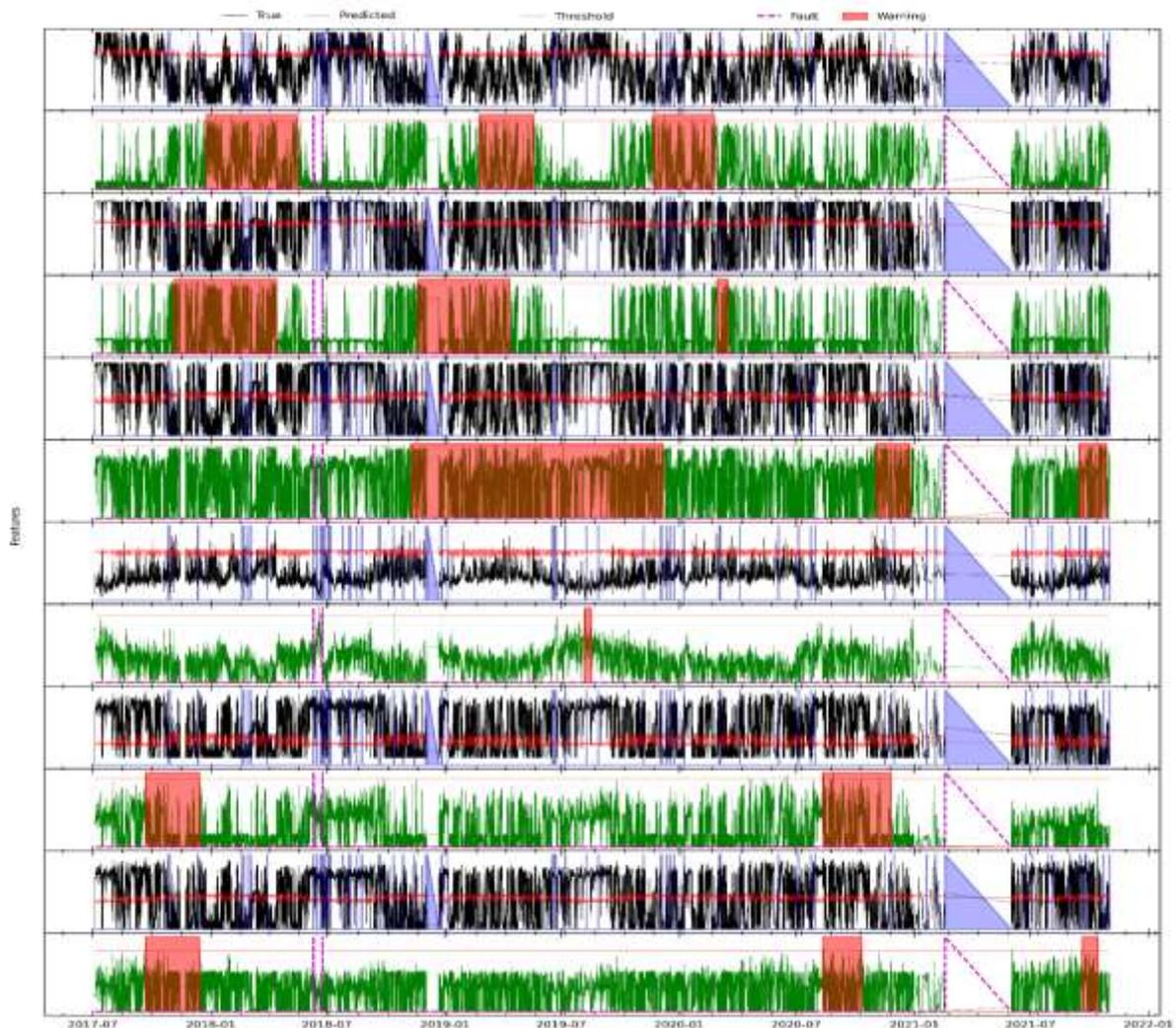


Figura 130. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.

En la **Figura 131** se destaca que el modelo advirtió la primera falla con seis meses de anticipación aproximadamente y también registró un FP y un FN.

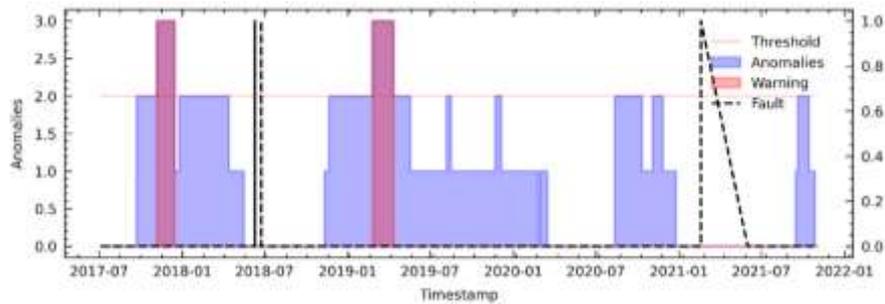


Figura 131. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.

C. TranAD utilizando el método PV18 en WT2

En la **Figura 132** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

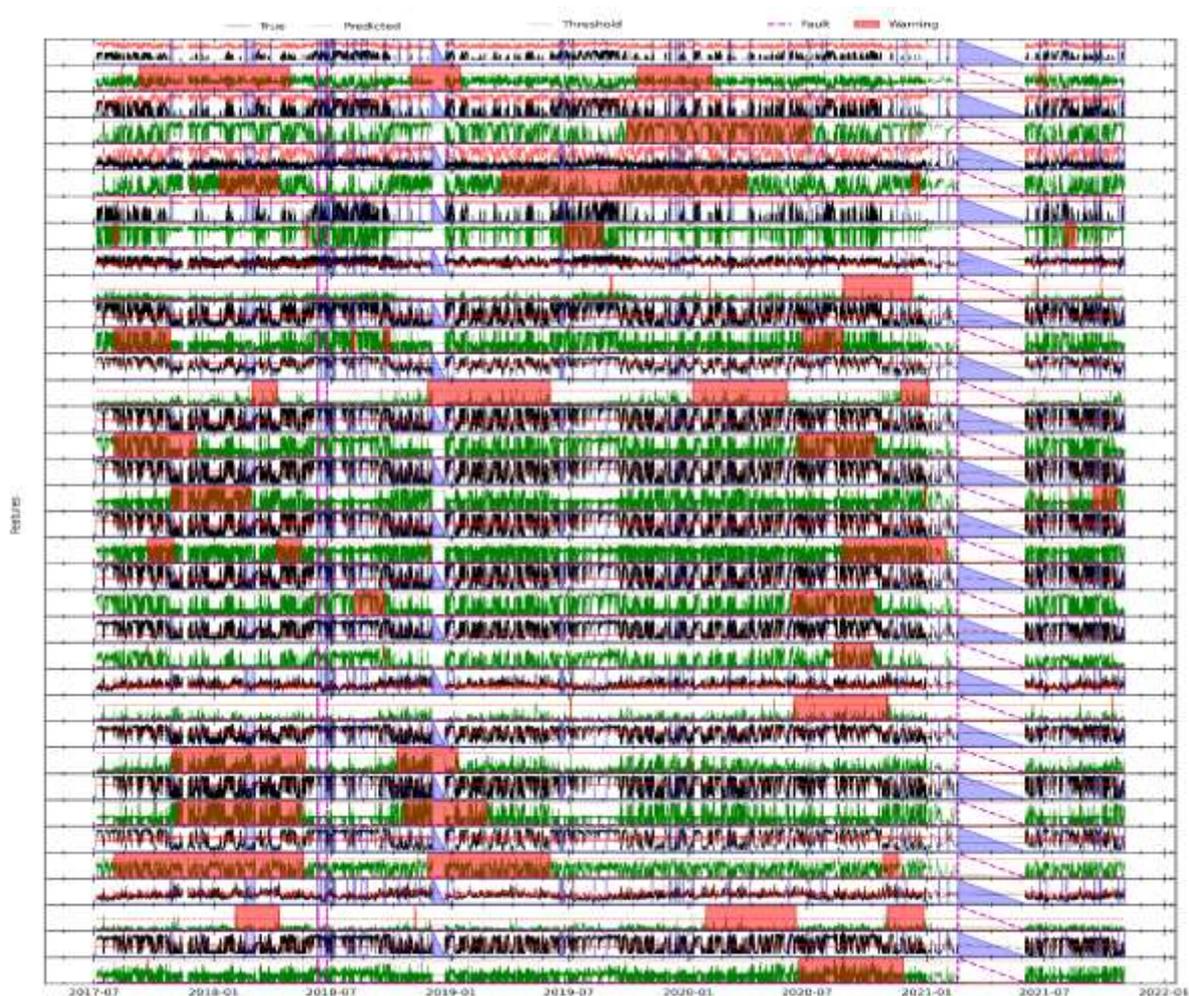


Figura 132. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.

En la **Figura 133** se muestra que el modelo advirtió las dos fallas en el módulo IGBT, con un tiempo de anticipación promedio aproximado de 6 meses.

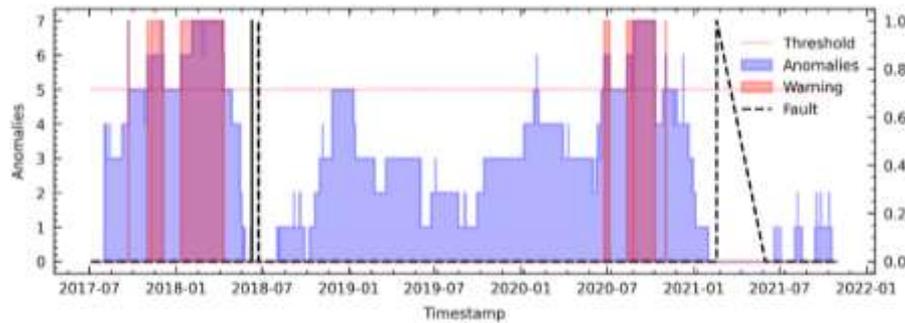


Figura 133. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.

D. TranAD utilizando el método LM18 en WT2

Para el análisis de LM18 en primer lugar, se debe analizar la **Figura 134**, donde se evalúa variable por variable.

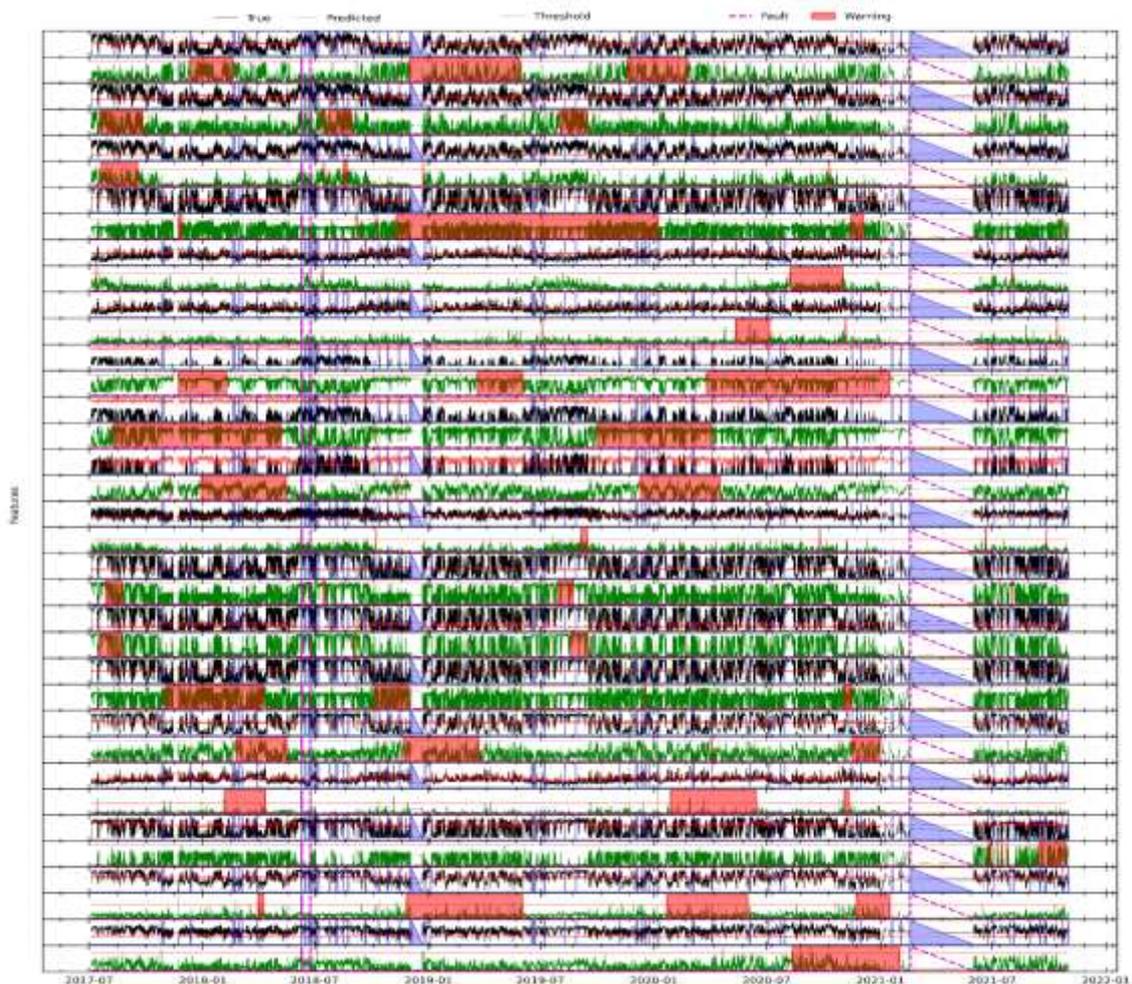


Figura 134. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT2.

En los resultados para LM18 se obtuvo la **Figura 135** donde se observa que el modelo advirtió las dos fallas con un promedio aproximado de 6 meses antes de la falla, y también se obtuvo tres FPs.

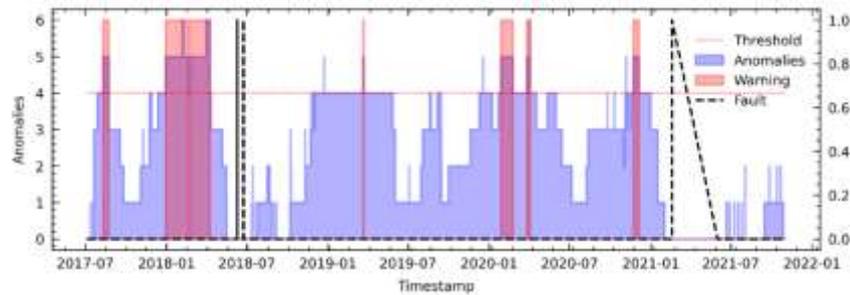


Figura 135. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT2.

XIII. Detección de fallas en WT3 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT3

En la **Figura 136** se analiza variable a variable la evaluación de fallas de primera etapa.

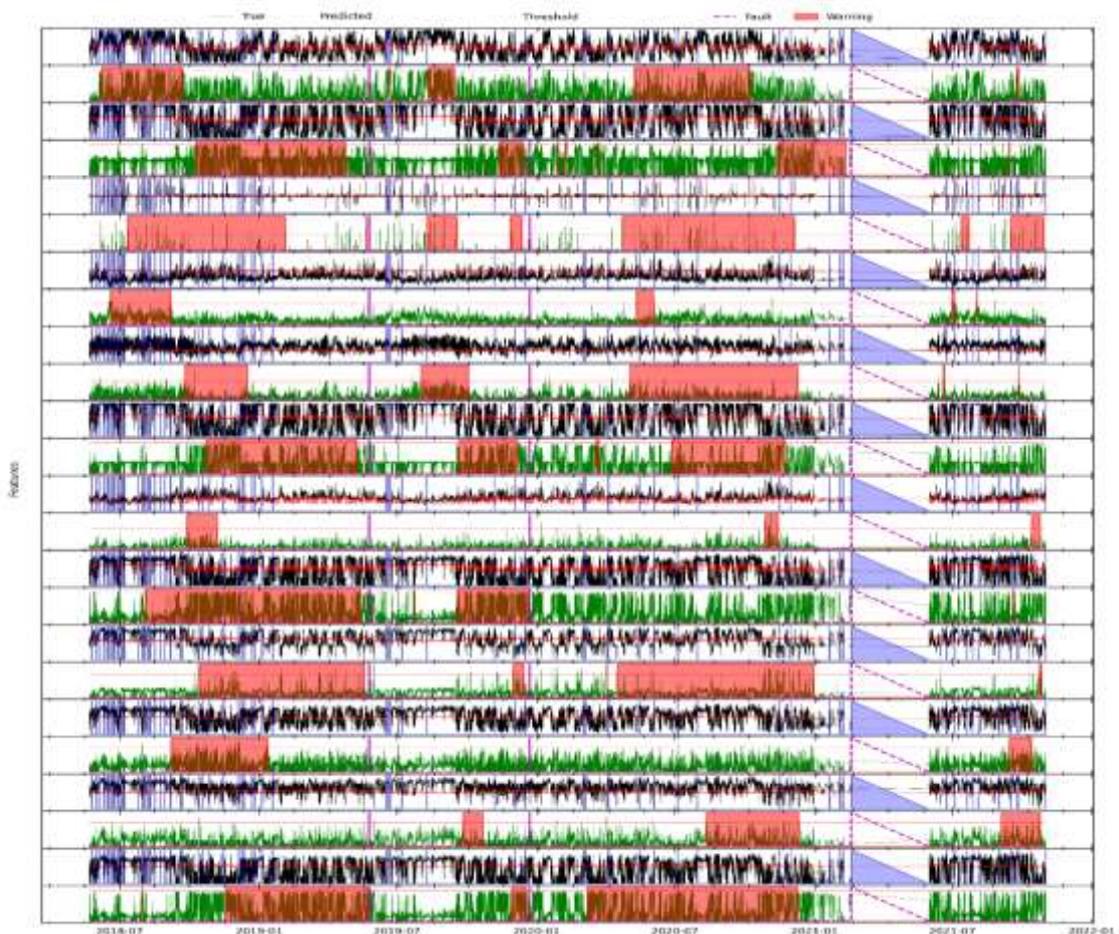


Figura 136. Resultados para RFE12 en el análisis de advertencias mediante el modelo TranAD para WT3.

En la **Figura 137** se destaca que el modelo advirtió las tres fallas en el módulo IGBT con un avance promedio de 6 meses aproximadamente, aunque también registro un FP al final del año 2021.

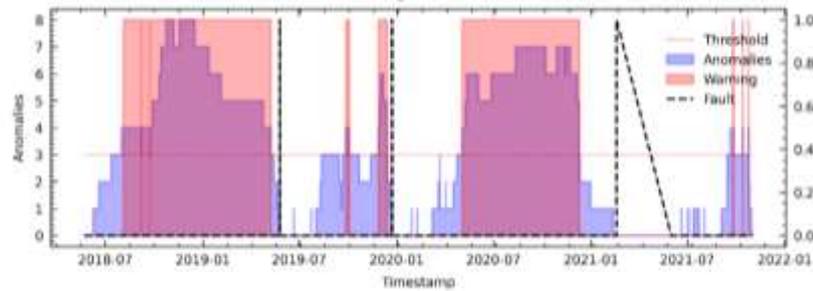


Figura 137. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.

B. TranAD utilizando el método C6 en WT3

En la **Figura 138** se analiza variable a variable la evaluación de fallas de primera etapa.

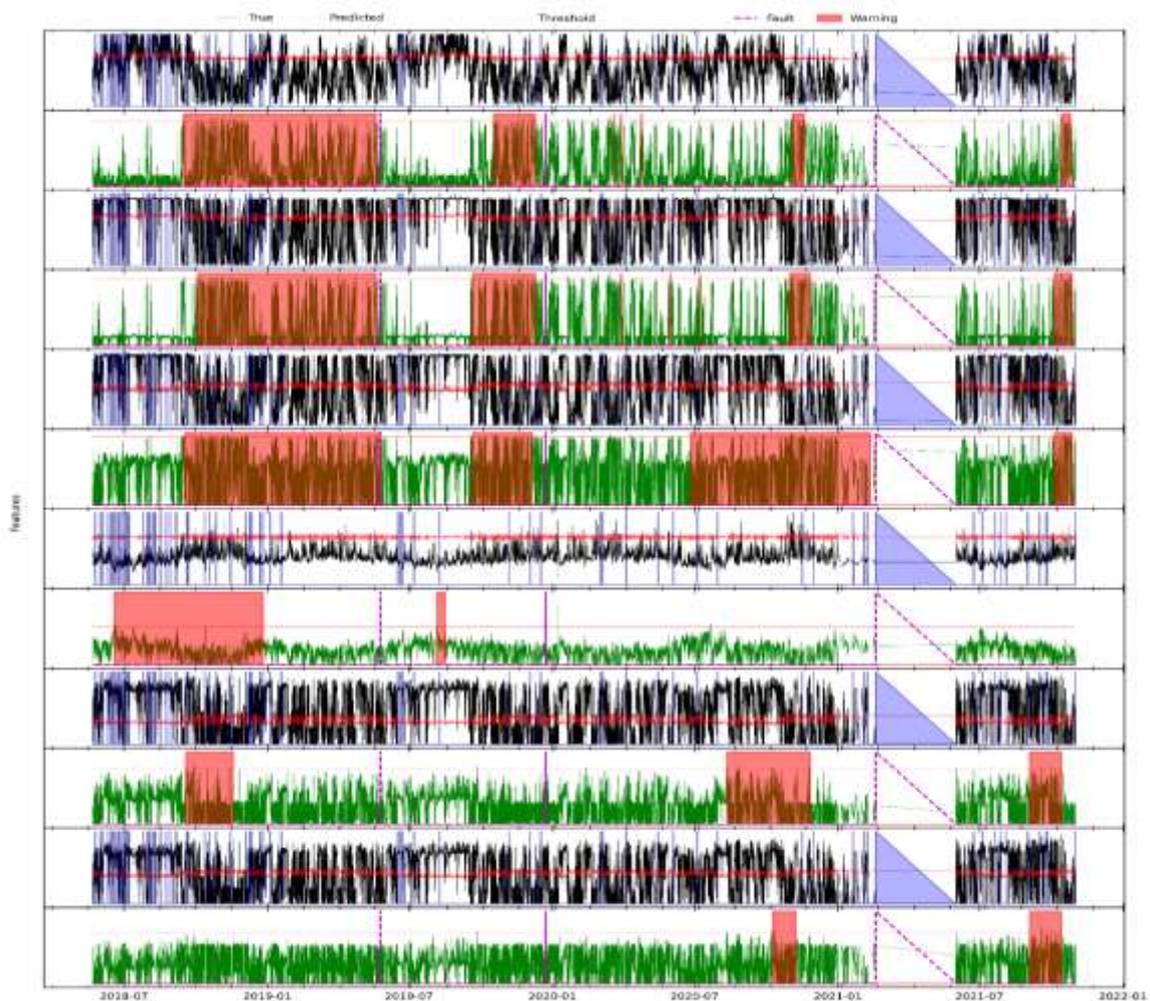


Figura 138. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT3.

En la **Figura 139** se destaca que el modelo advirtió las tres fallas en el módulo IGBT con un avance aproximado de 5 meses, aunque también registró un FP.

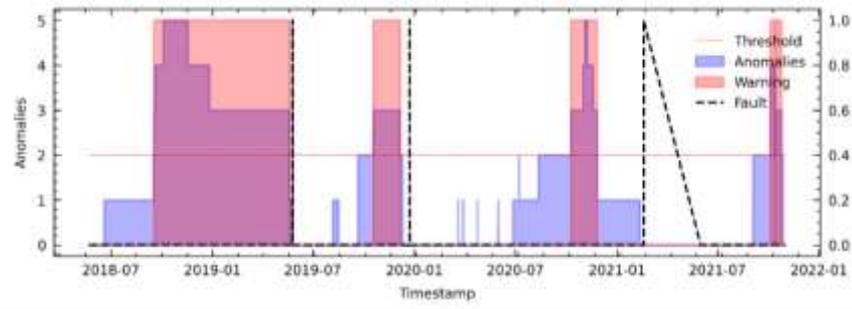


Figura 139. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.

C. TranAD utilizando el método PV18 en WT3

En la **Figura 140** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

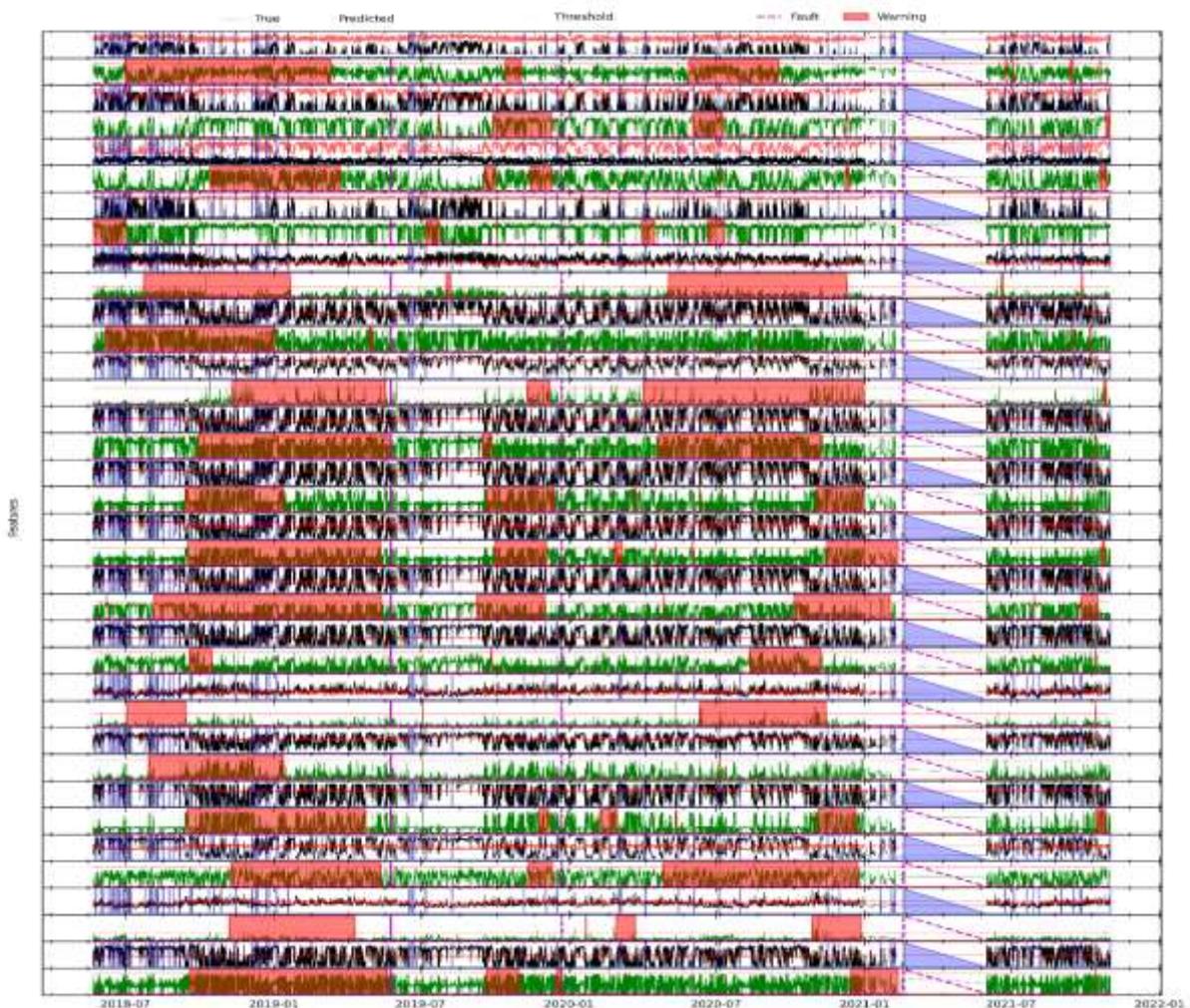


Figura 140. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT3.

En la **Figura 141** se muestra que el modelo pudo advertir las tres fallas en el módulo IGBT con un promedio aproximado de 7 meses.

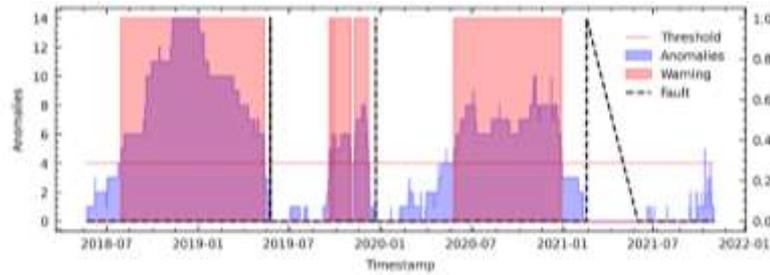


Figura 141. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.

D. TranAD utilizando el método LM18 en WT3

Para el análisis de LM18, en primer lugar, se debe observar la **Figura 142**, donde se analiza variable por variable la evaluación de fallas.

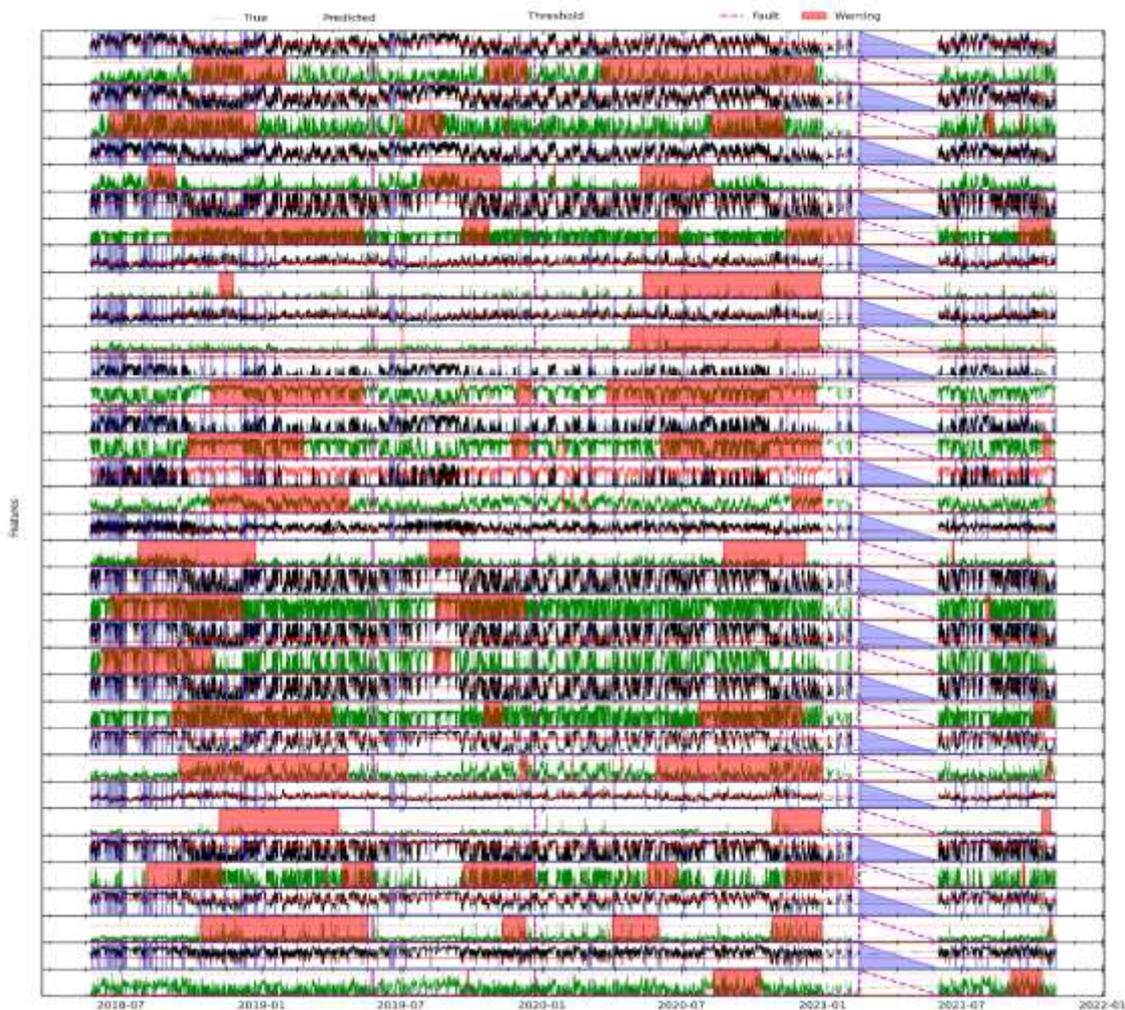


Figura 142. Resultados para LM18 en el análisis de advertencias variable a variable en relación a las fallas en el módulo IGBT mediante el modelo TranAD para WT3.

En los resultados para LM18 se obtuvo la **Figura 143**, donde se observa que el modelo advirtió las fallas con un promedio aproximado de 6 meses antes de su ocurrencia, así mismo registro un FP a finales de 2021.

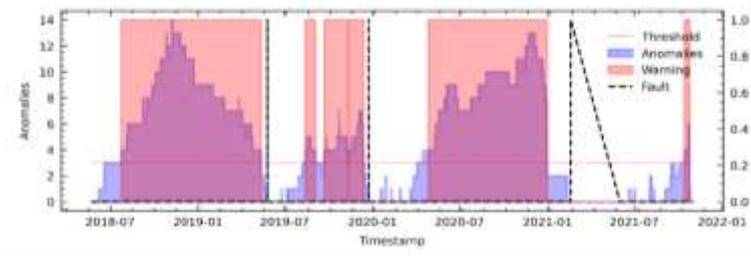


Figura 143. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT3.

XIV. Detección de fallas en WT4 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT4

En la **Figura 144** se analiza variable a variable la evaluación de fallas de primera etapa.

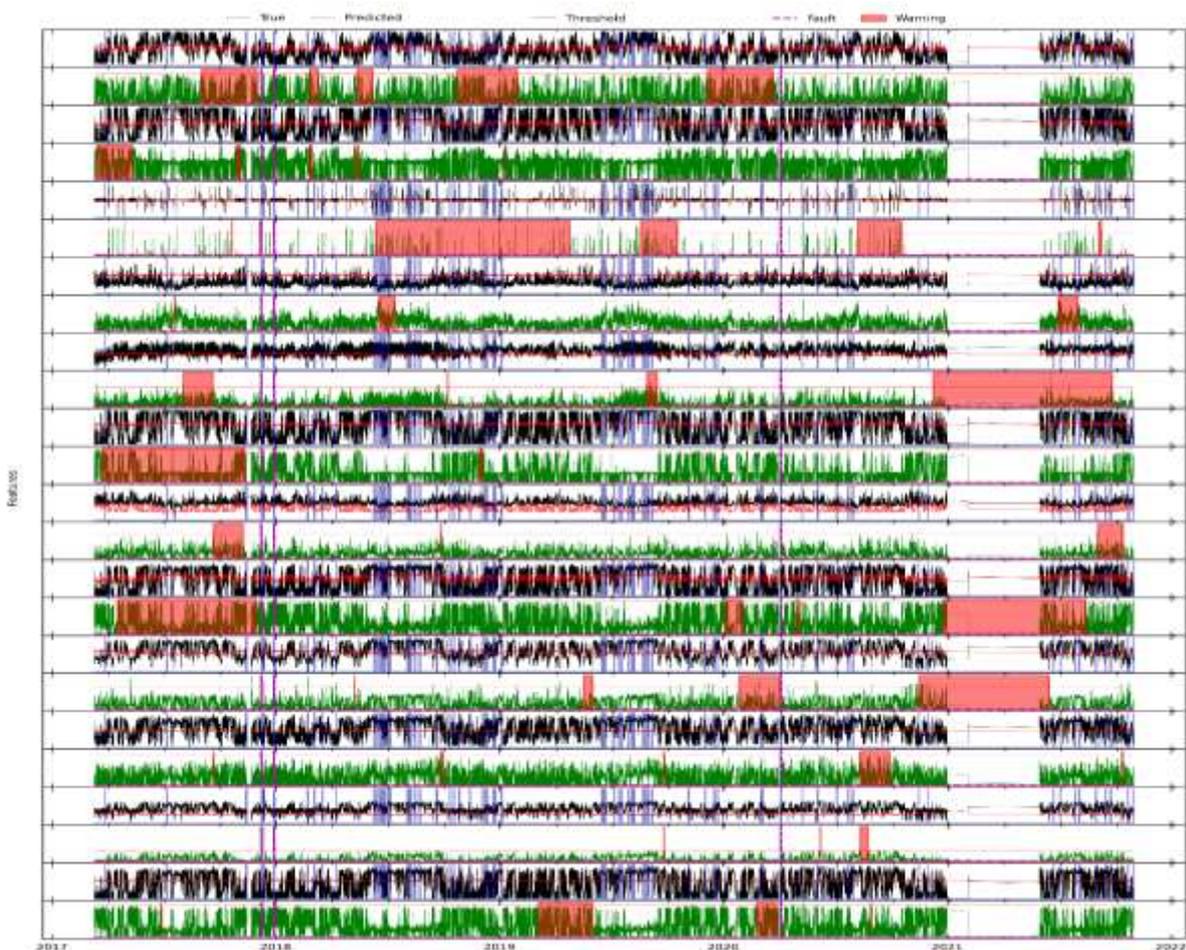


Figura 144. Resultados para RFE12 en el análisis de advertencias variable mediante el modelo TranAD para WT4.

En la **Figura 145** se destaca que el modelo advirtió una falla de las dos que se encuentran registradas en el archivo de O&M.

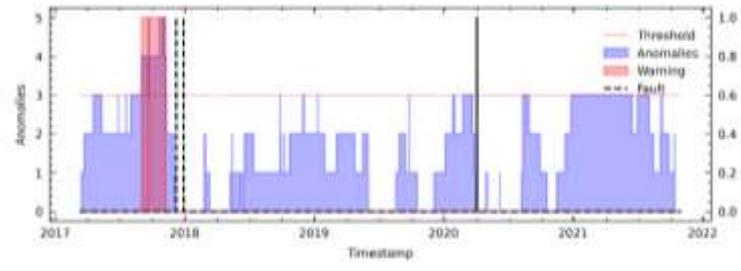


Figura 145. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.

B. TranAD utilizando el método C6 en WT4

En la **Figura 146** se analiza variable a variable la evaluación de fallas de primera etapa.

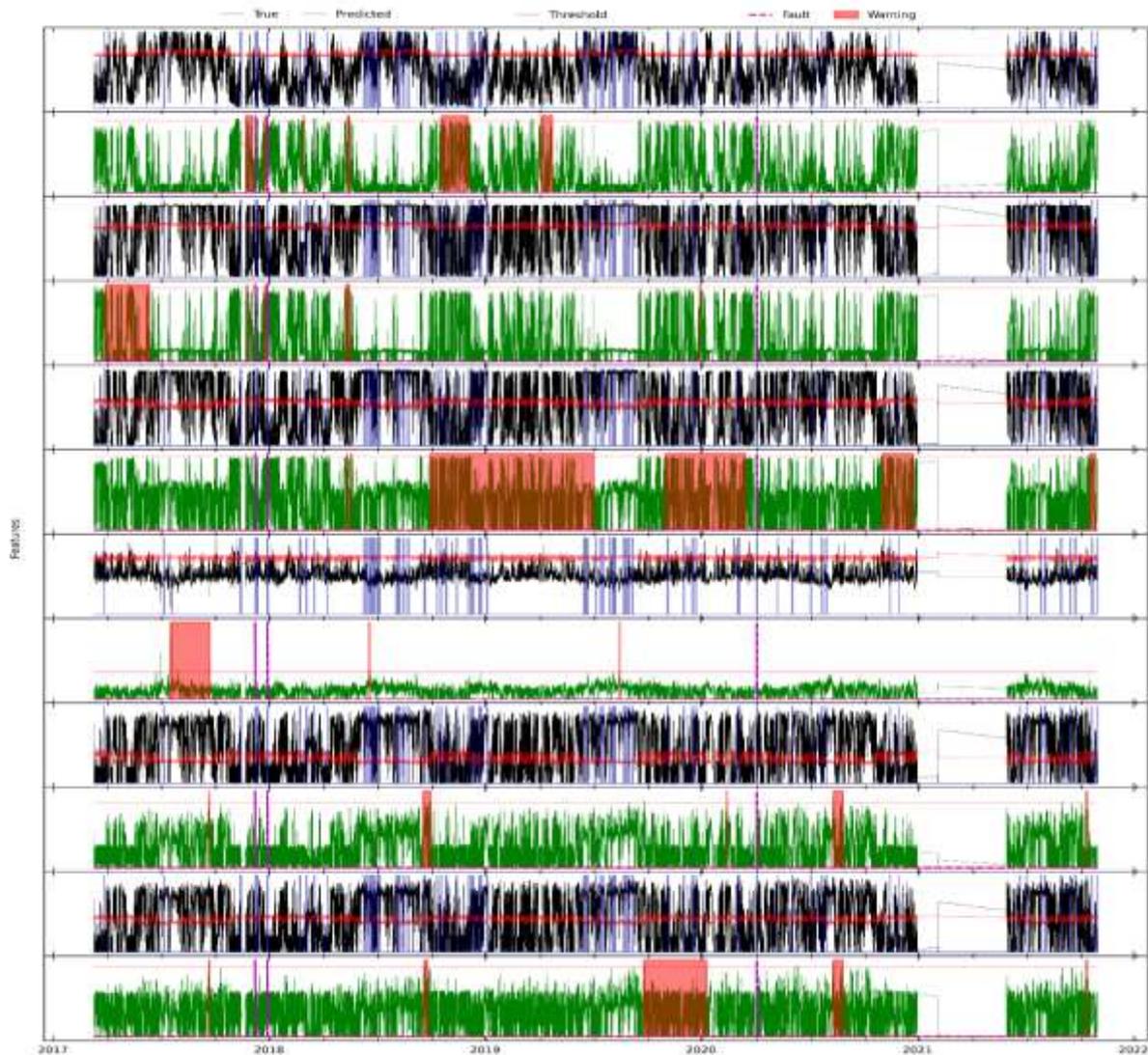


Figura 146. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT4.

En la **Figura 147** se destaca que el modelo solo advirtió una de las dos fallas en el módulo IGBT.

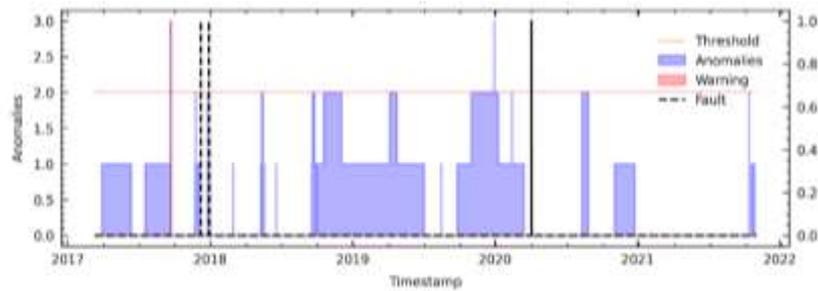


Figura 147. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.

C. TranAD utilizando el método PV18 en WT4

En la **Figura 148** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

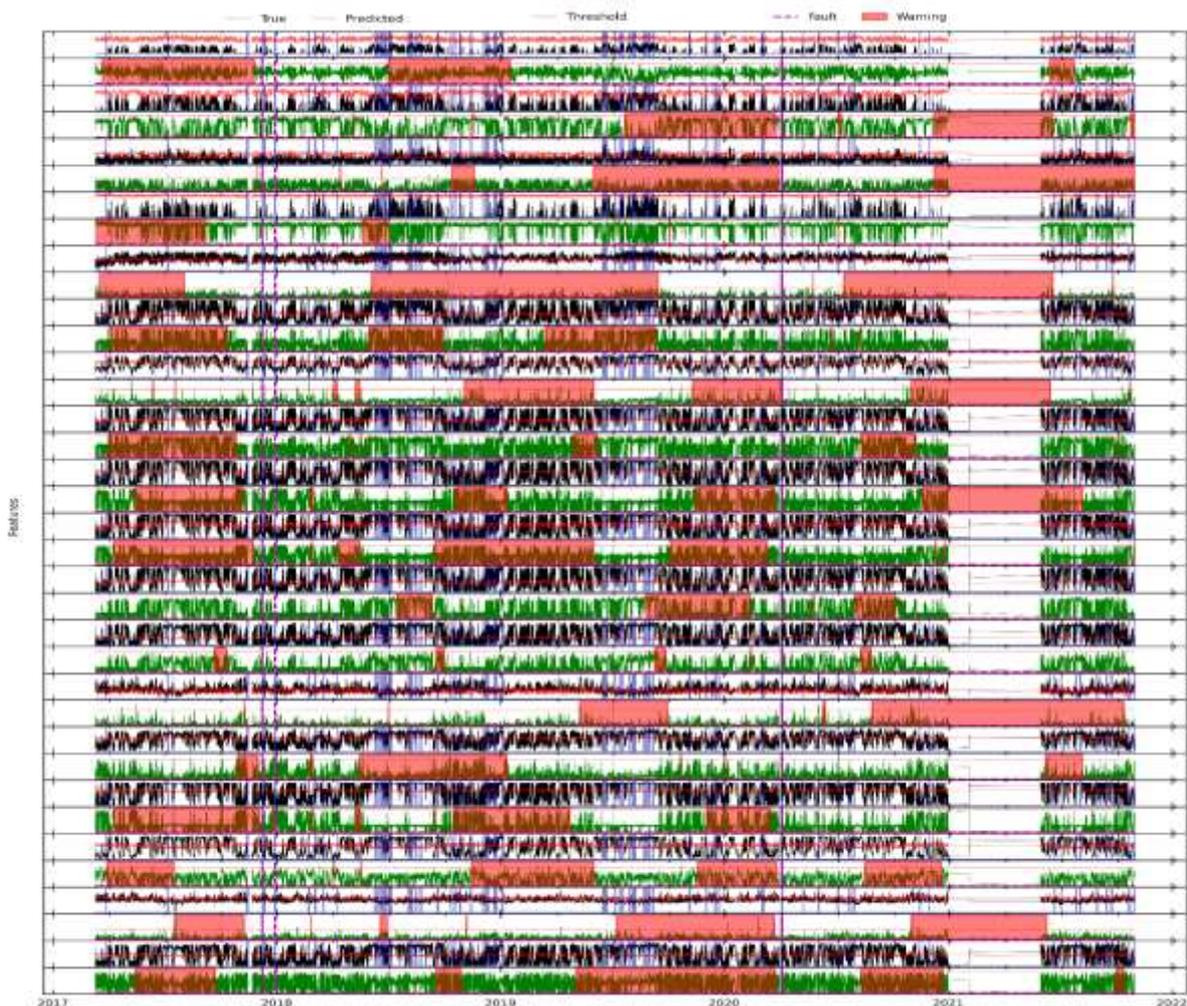


Figura 148. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT4.

En la **Figura 149** se muestra que el modelo si logró advertir las dos fallas en el módulo IGBT con un avance aproximado de 4 meses de anticipación, pero también dio un FN.

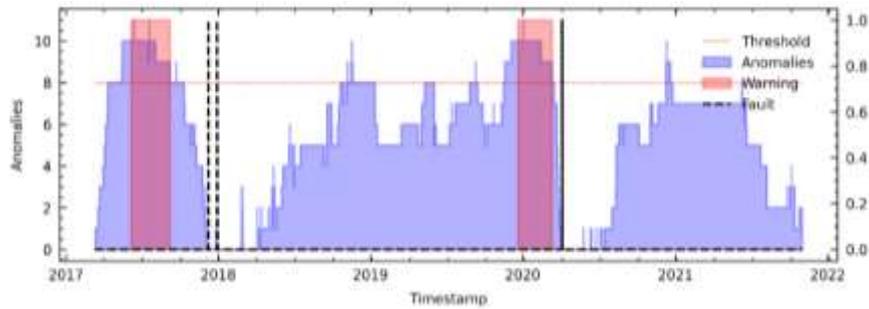


Figura 149. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.

D. TranAD utilizando el método LM18 en WT4

Para el análisis de LM18 variable a variable como se muestra en la **Figura 150**.

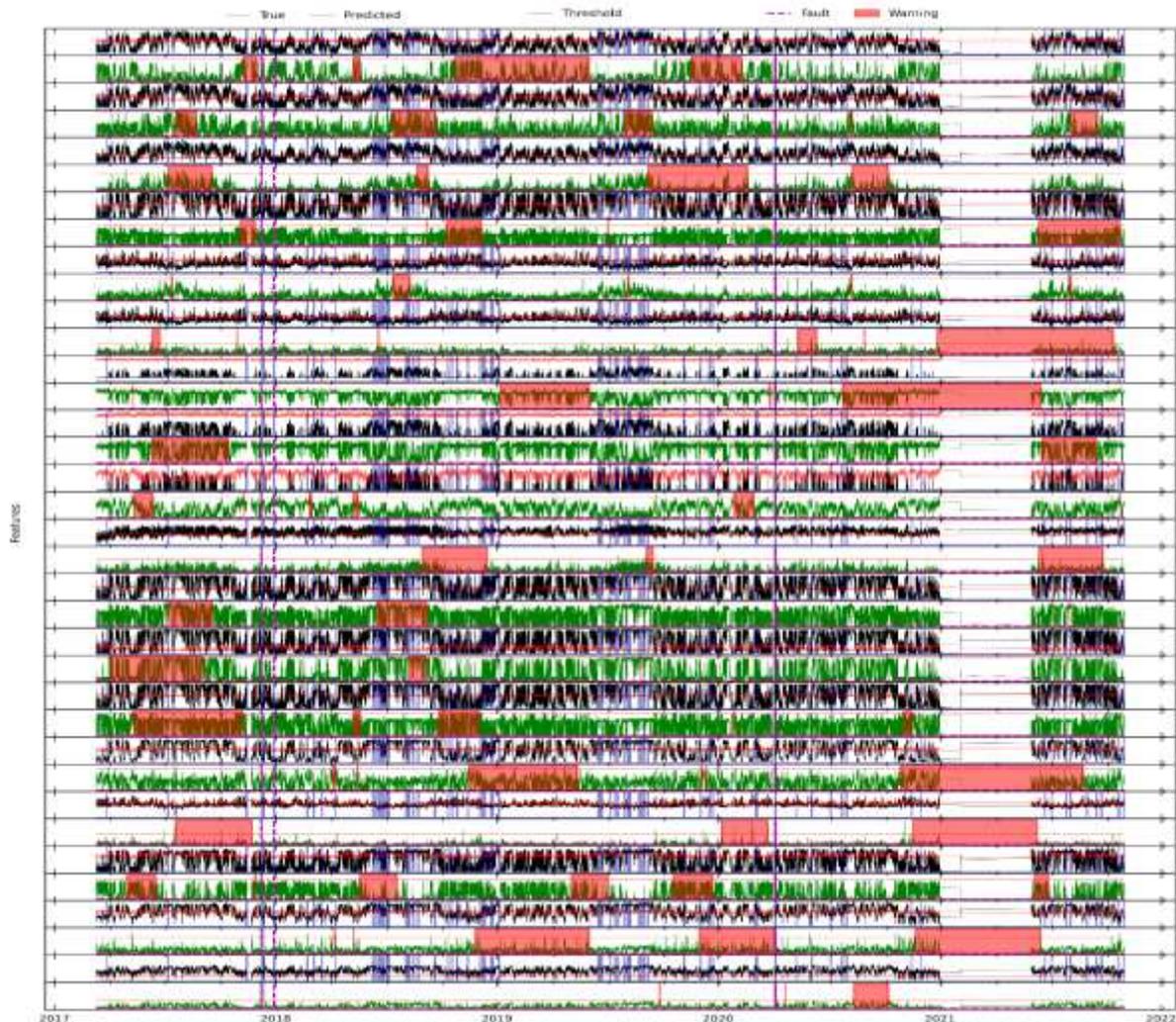


Figura 150. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT4.

En los resultados para LM18 se obtuvo la **Figura 151**, donde se observa que el modelo advirtió las dos fallas con un avance aproximado de 3 meses y también se registra tres FPs.

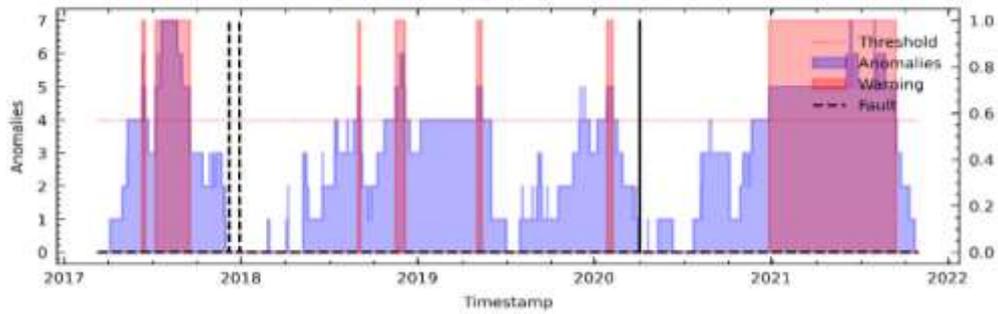


Figura 151. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT4.

XV. Detección de fallas en WT5 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT5

En la **Figura 152** se analiza variable a variable la evaluación de fallas de primera etapa.

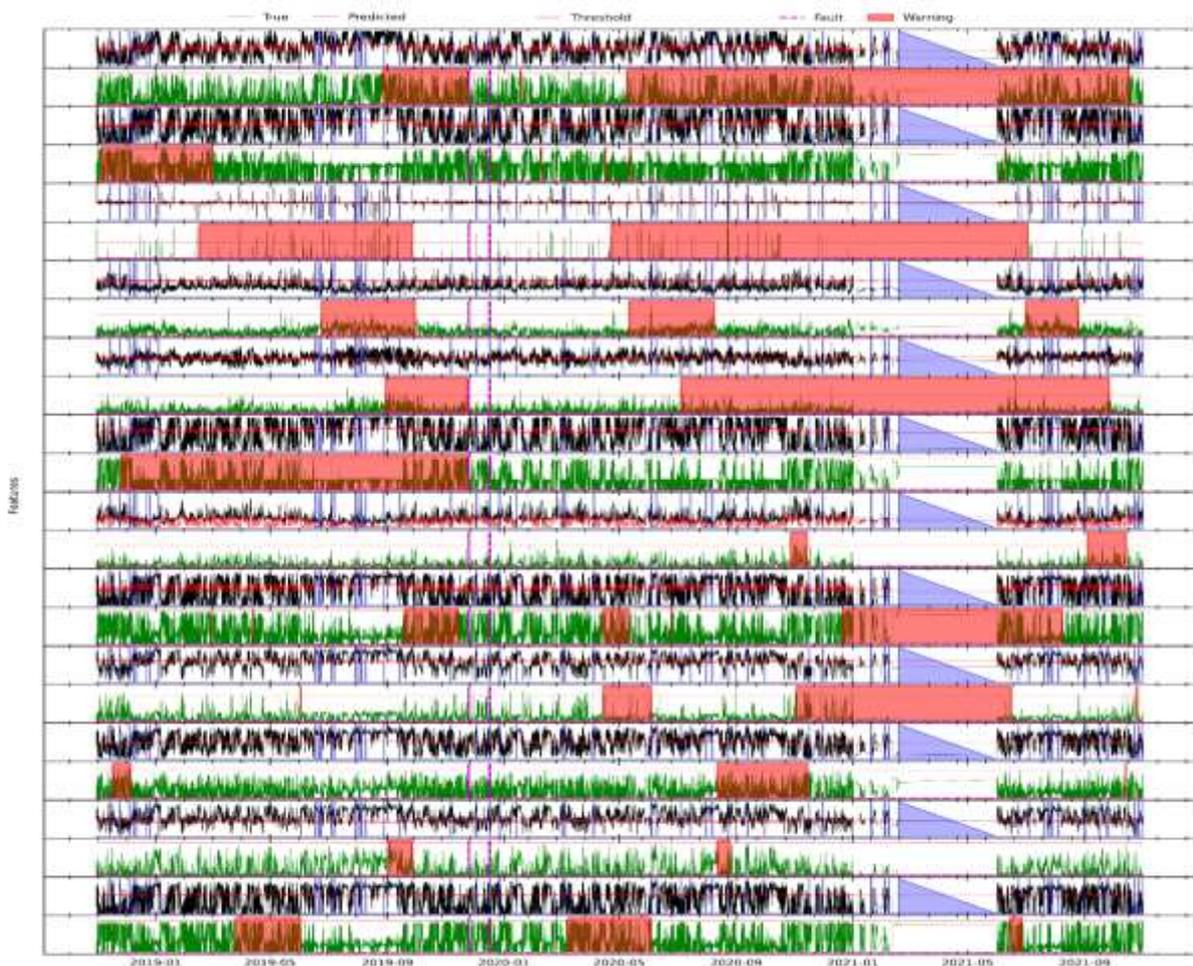


Figura 152. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para la WT5.

En la **Figura 153** se destaca que el modelo advirtió la falla registrada en el módulo IGBT con un avance aproximado de 3 meses.

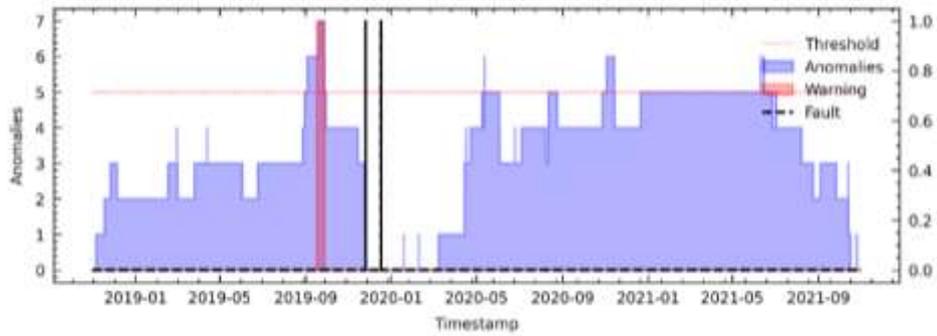


Figura 153. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.

B. TranAD utilizando el método C6 en WT5

En la **Figura 154** se analiza variable a variable la evaluación de fallas de primera etapa.

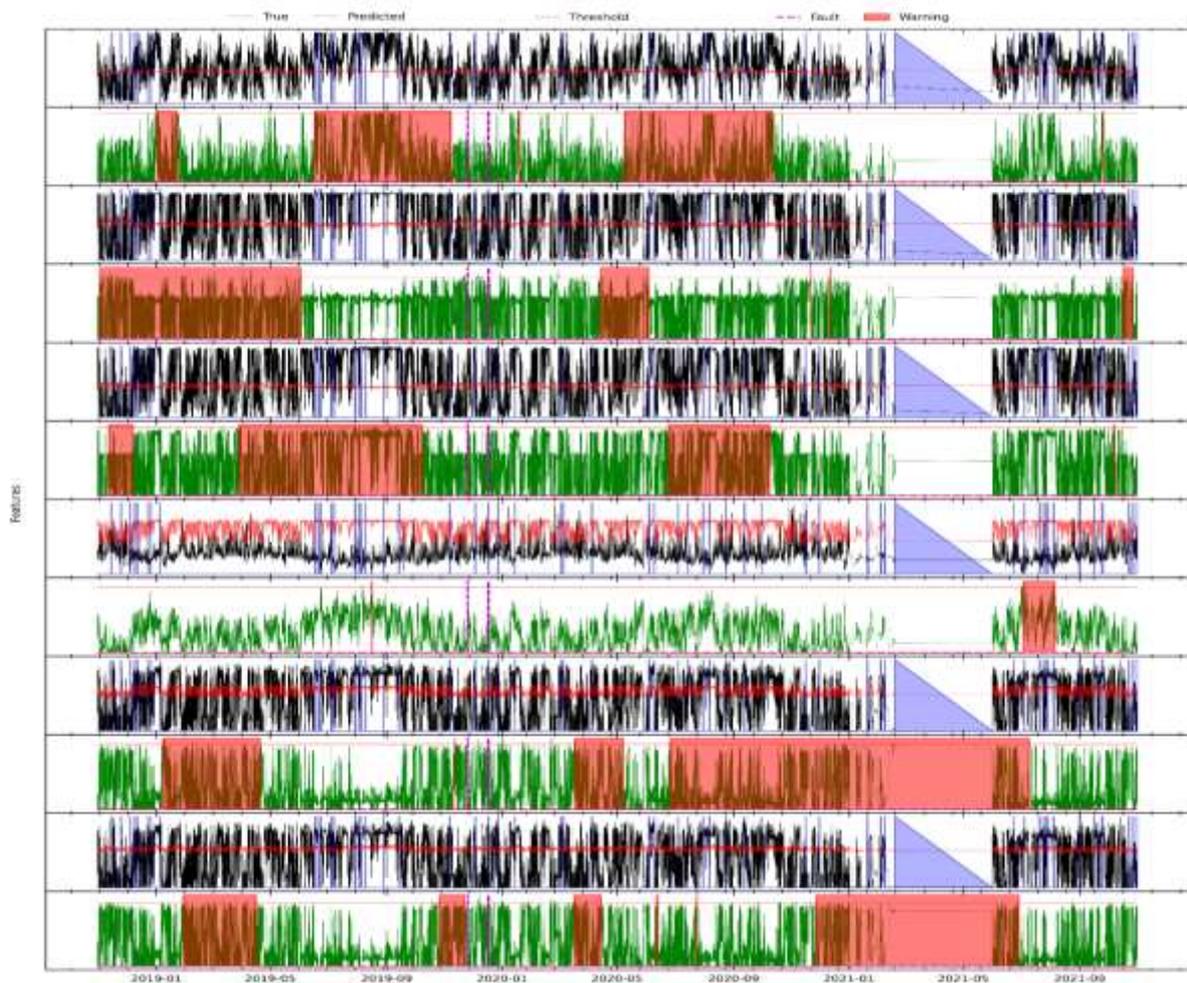


Figura 154. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT5.

En la **Figura 155** en la que se destaca que el modelo advirtió la falla del módulo IGBT con un avance aproximado de 8 meses y registró un FP.

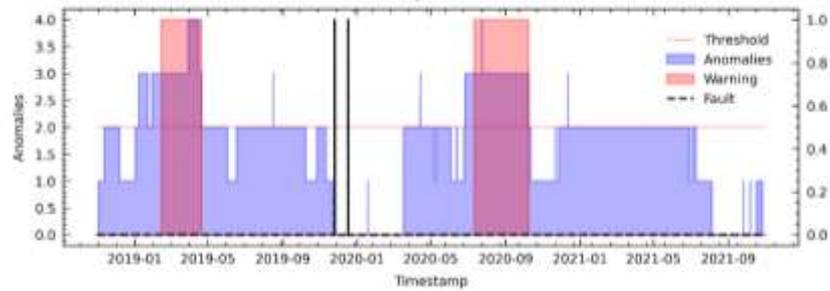


Figura 155. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.

C. TranAD utilizando el método PV18 en WT5

En la **Figura 156** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

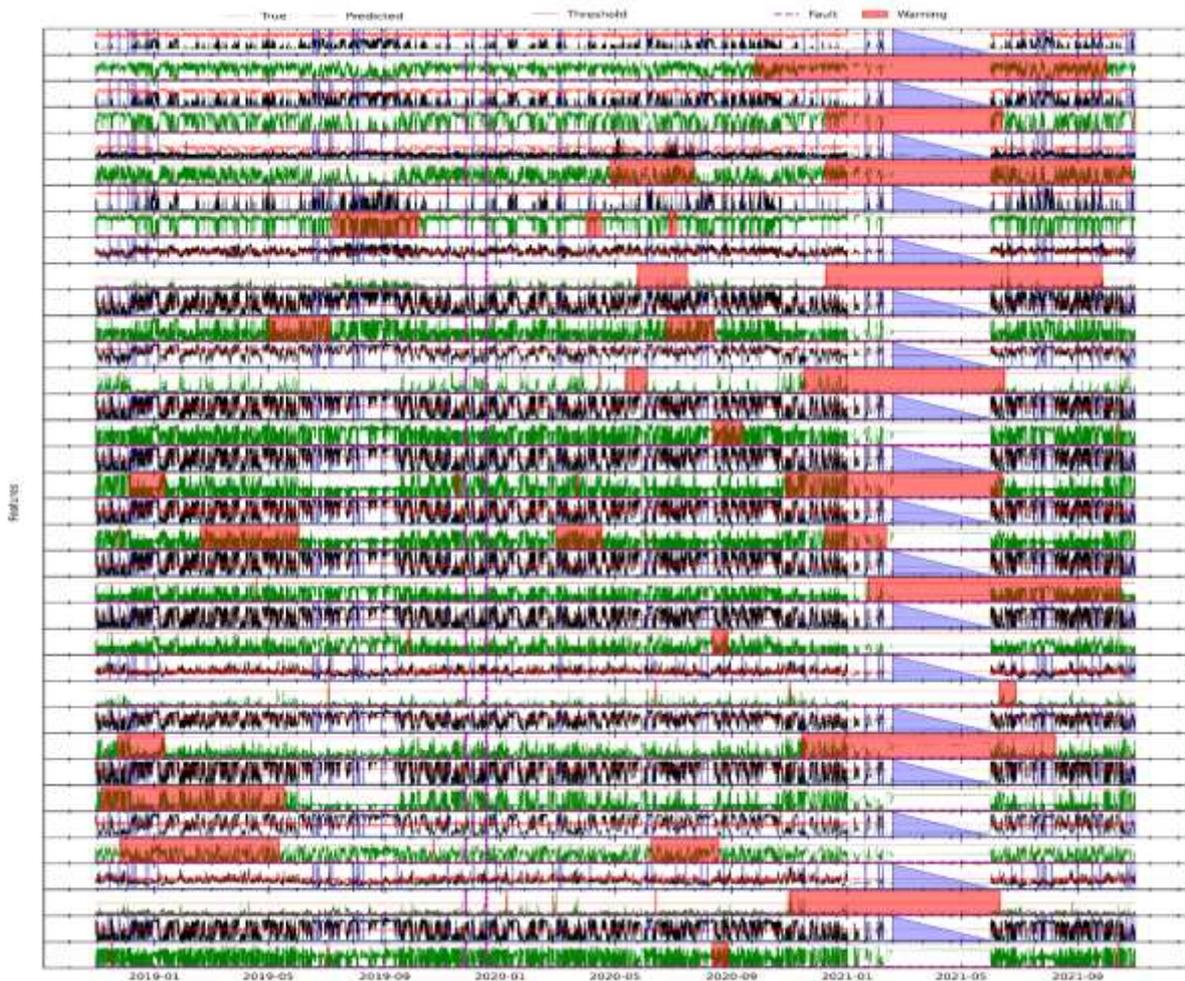


Figura 156. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT5.

En la **Figura 157** se muestra que el modelo no logró advertir las dos fallas en el módulo IGBT, pero también dio un FN.

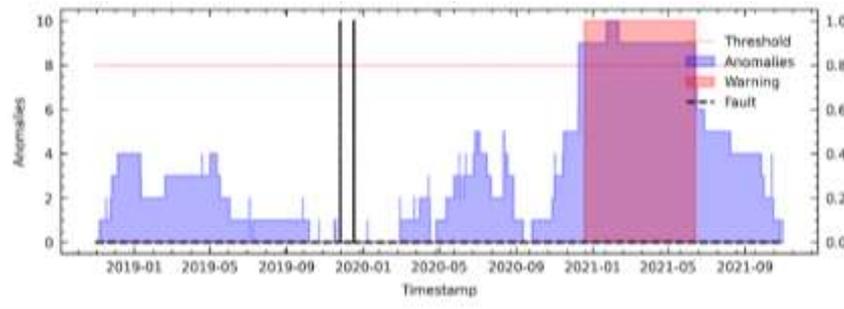


Figura 157. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.

D. TranAD utilizando el método LM18 en WT5

Para el análisis de LM18, en primer lugar, se debe observar la evaluación de primera etapa como se muestra en la **Figura 158**.

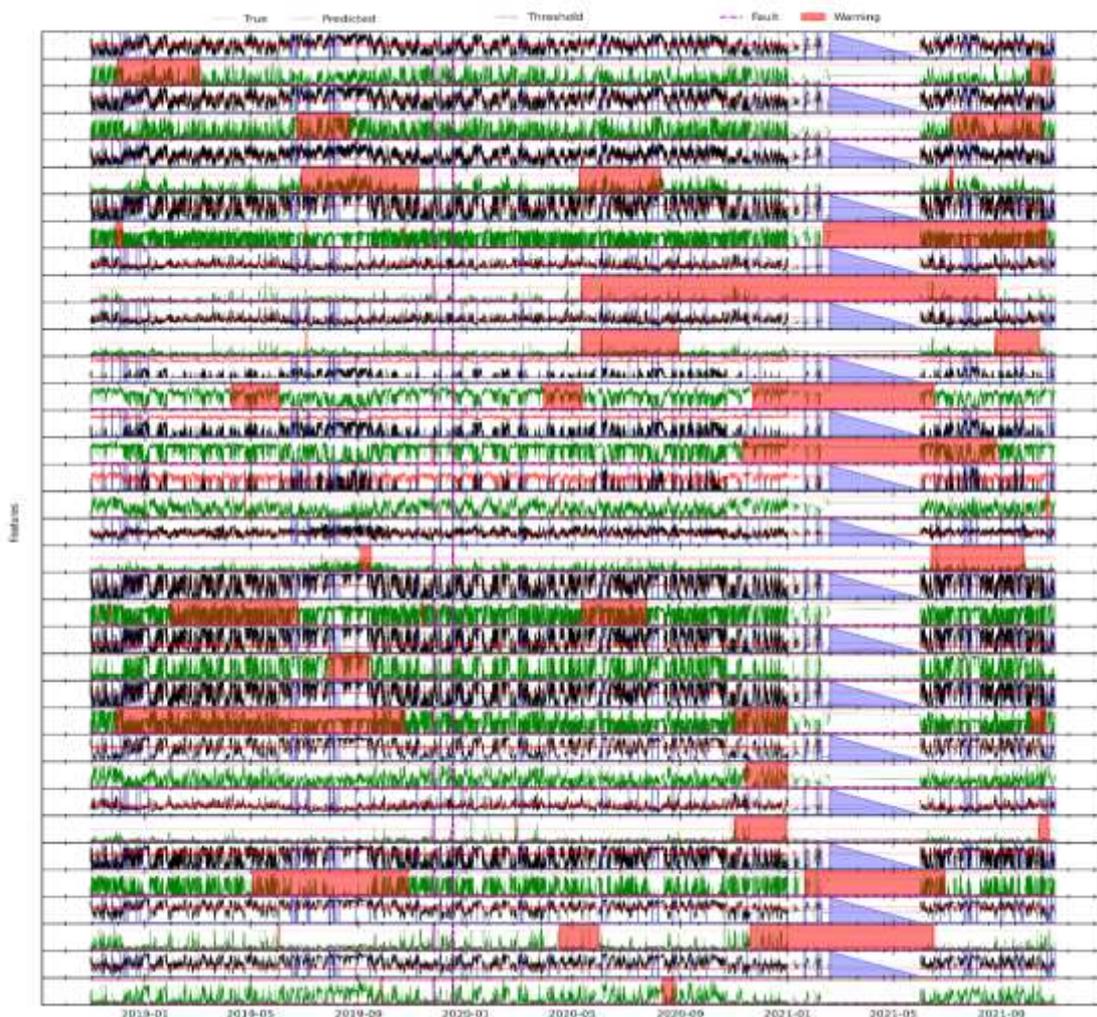


Figura 158. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT5.

En los resultados para LM18 se obtuvo la **Figura 159**, donde se observa que el modelo no logró advertir la falla registrada en el archivo de O&M y registró un FN.

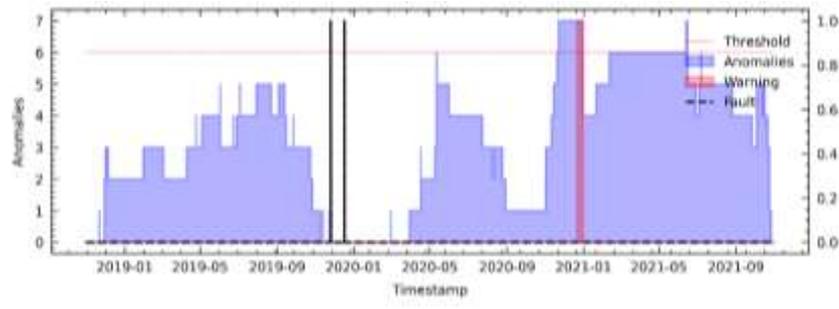


Figura 159. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT5.

XVI. Detección de fallas en WT6 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT6

En la **Figura 160** se analiza variable a variable la evaluación de fallas de primera etapa.

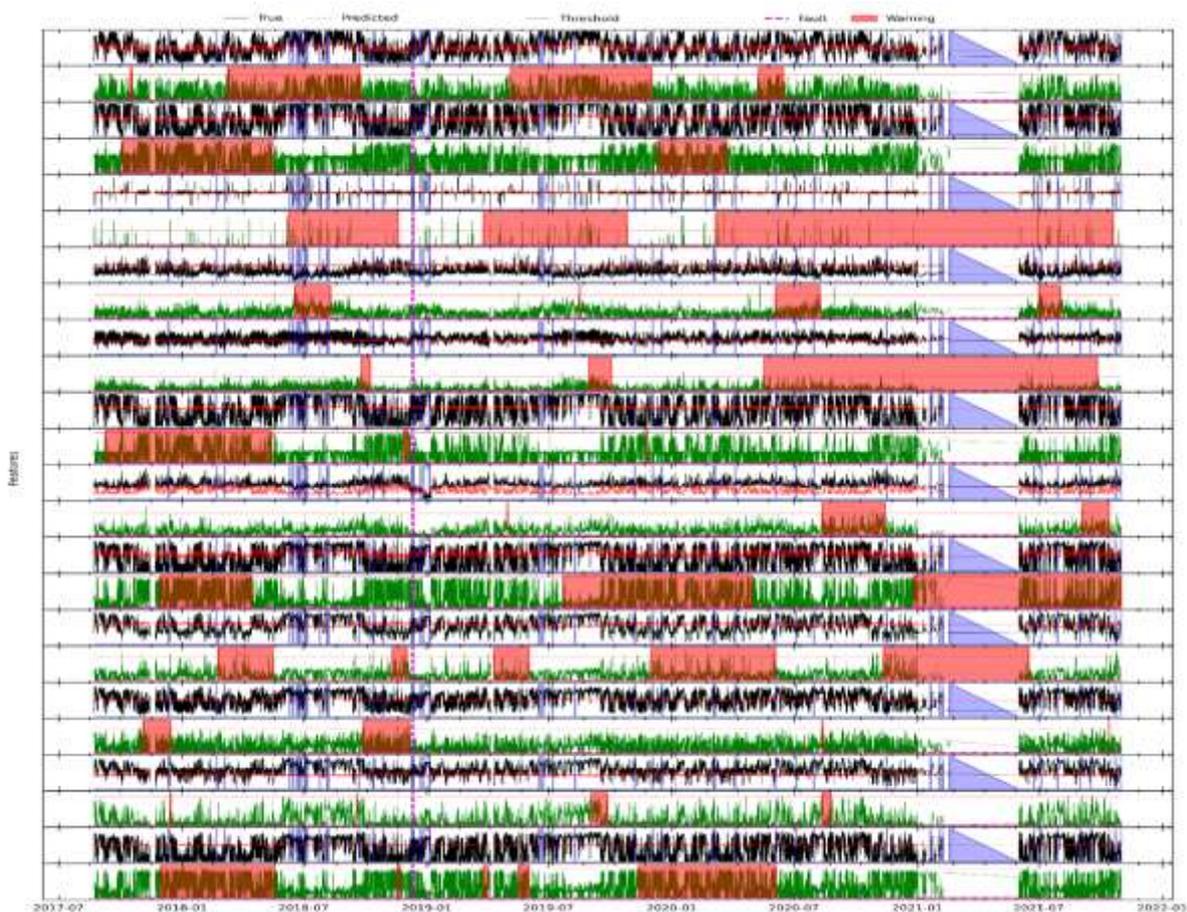


Figura 160. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.

En la **Figura 161** se destaca que el modelo advirtió una falla con un avance aproximado de 8 meses de anticipación y también registró un FP.

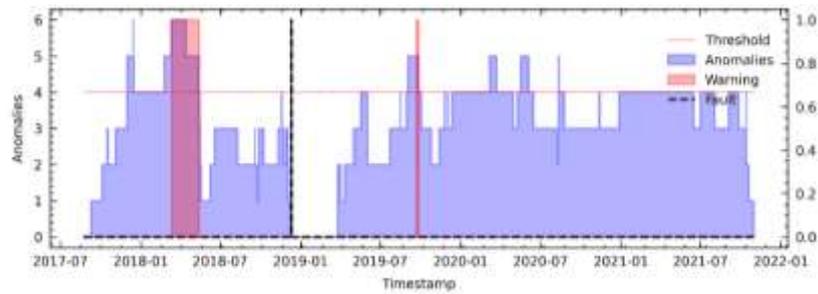


Figura 161. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.

B. TranAD utilizando el método C6 en WT6

En la **Figura 162** se analiza variable a variable la evaluación de fallas de primera etapa.

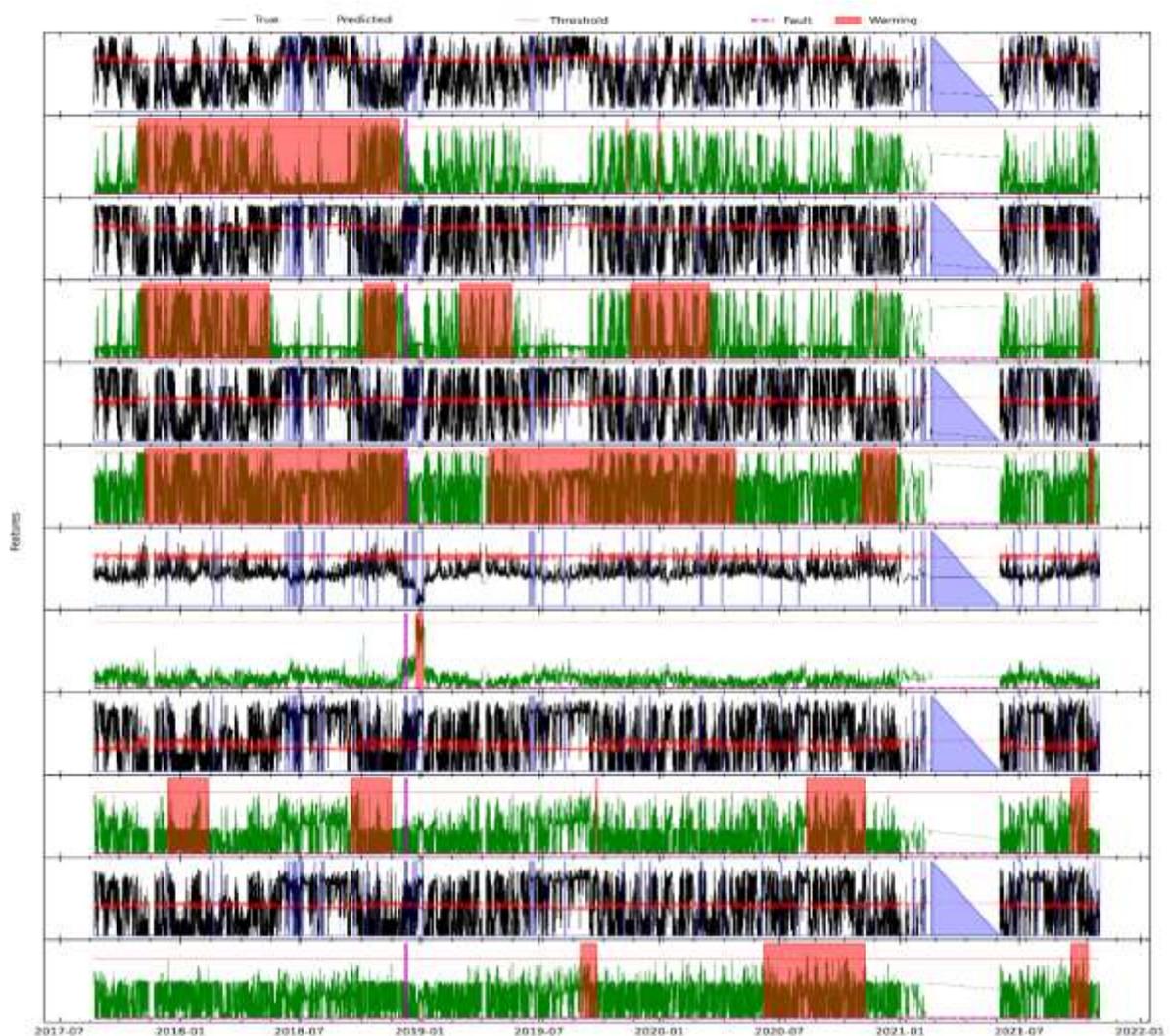


Figura 162. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.

En la **Figura 163** se destaca que el modelo advirtió la falla con un avance aproximado de 2 meses y registró un FP.

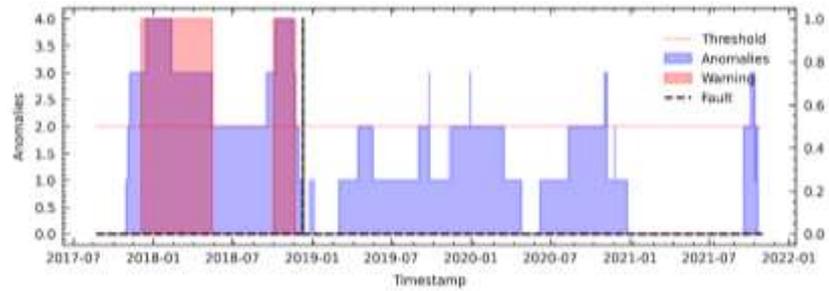


Figura 163. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.

C. TranAD utilizando el método PV18 en WT6

En la **Figura 164** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

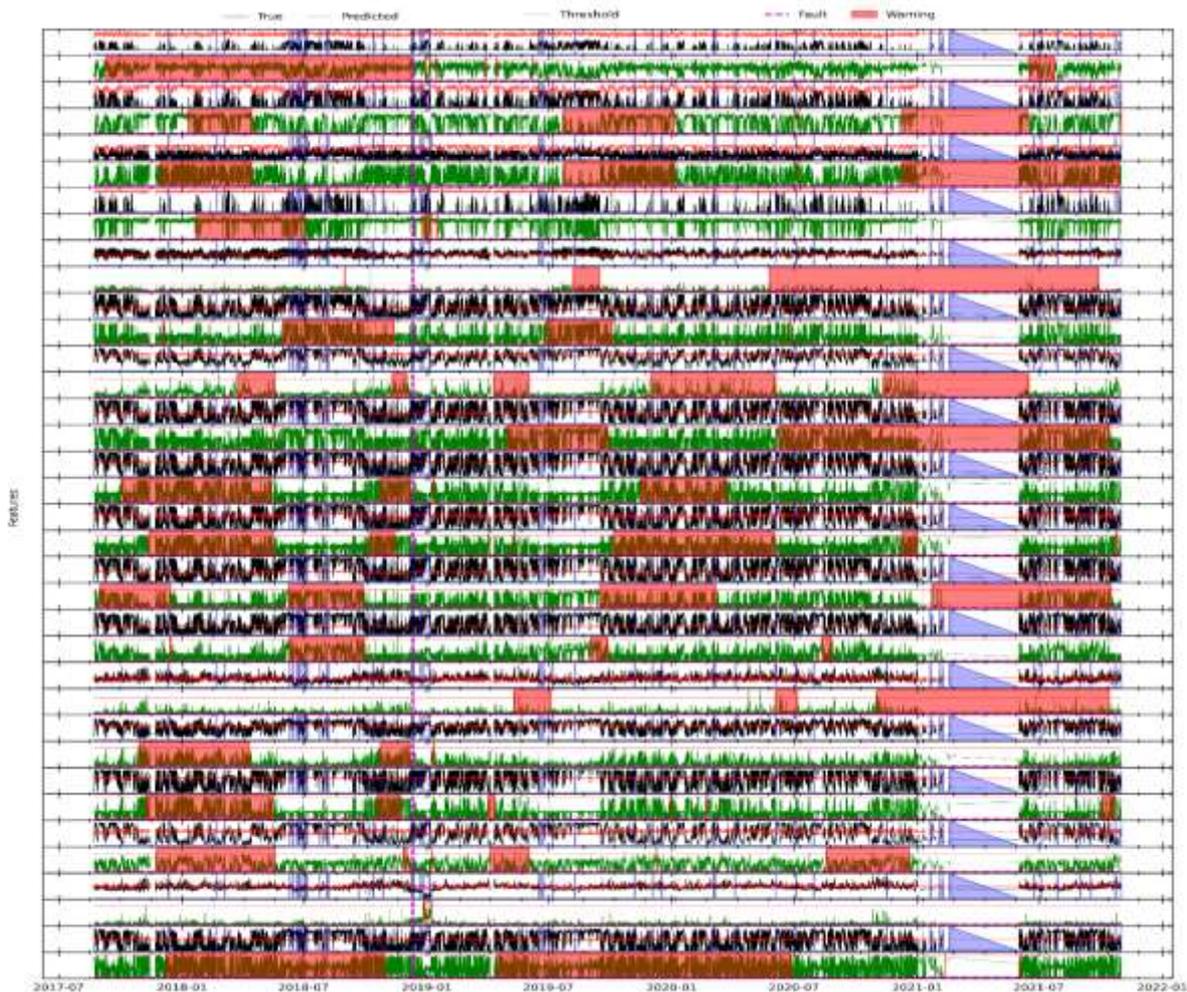


Figura 164. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.

En la **Figura 165** se muestra que el modelo si logró advertir la falla en el módulo IGBT con un avance aproximado 10 meses antes de que ocurriese.

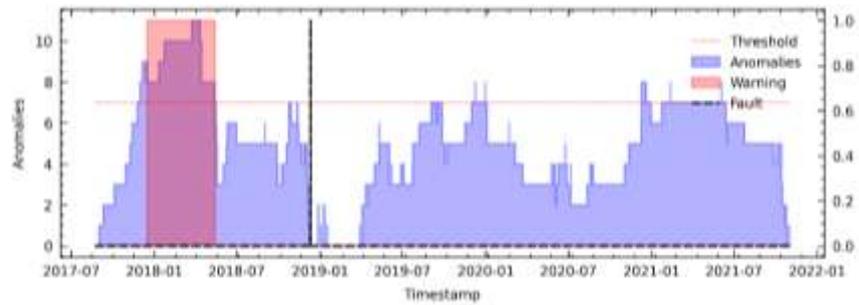


Figura 165. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.

D. TranAD utilizando el método LM18 en WT6

Para el análisis de LM18 se analiza la evaluación de primera etapa como se muestra en la **Figura 166**.

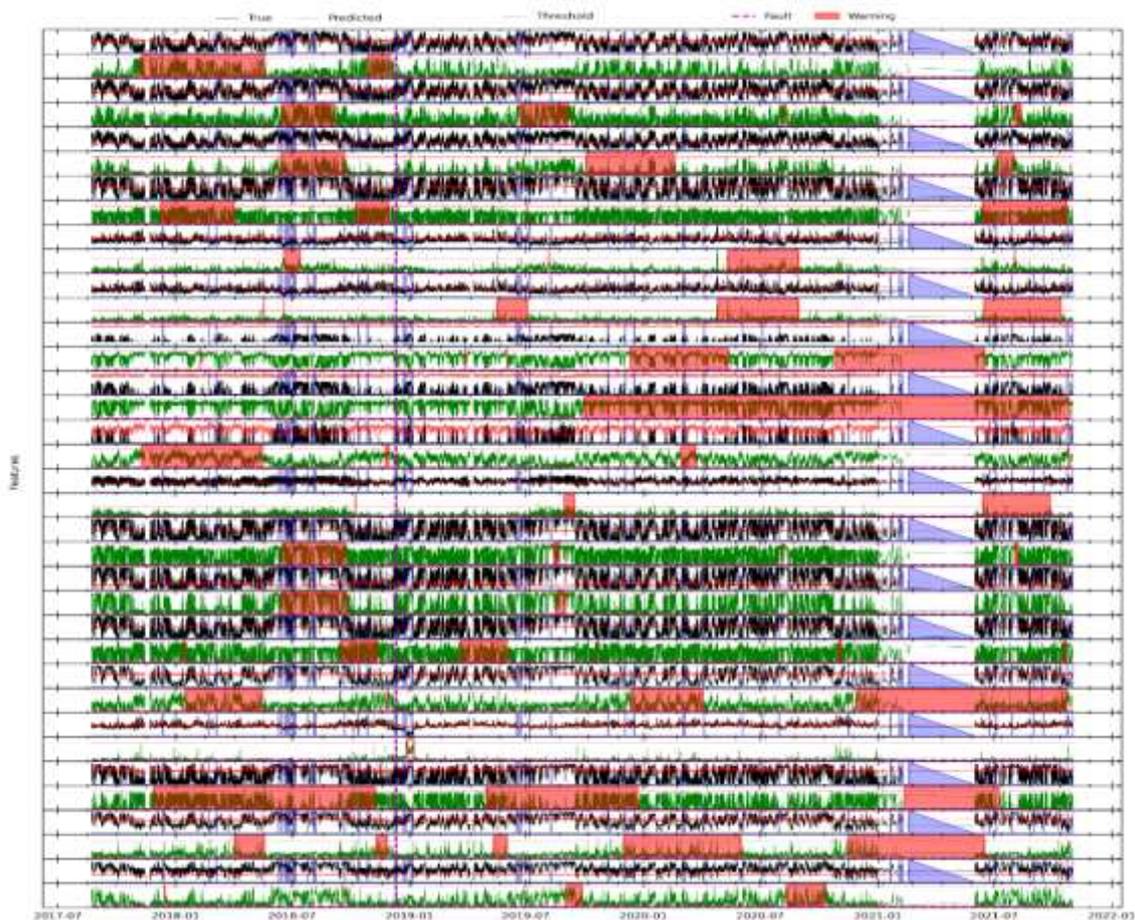


Figura 166. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT6.

En los resultados para LM18 se obtuvo la **Figura 167**, donde se observa que el modelo advirtió la falla del módulo IGBT con aproximadamente 6 meses de anticipación y también registró 2 FPs.

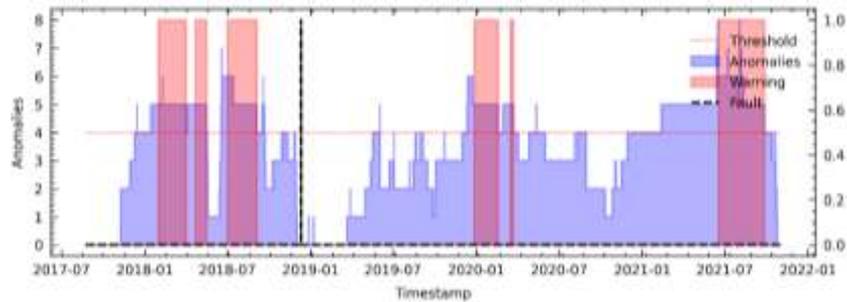


Figura 167. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT6.

XVII. Detección de fallas en WT7 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT7

En la **Figura 168** se analiza variable a variable la evaluación de fallas de primera etapa.

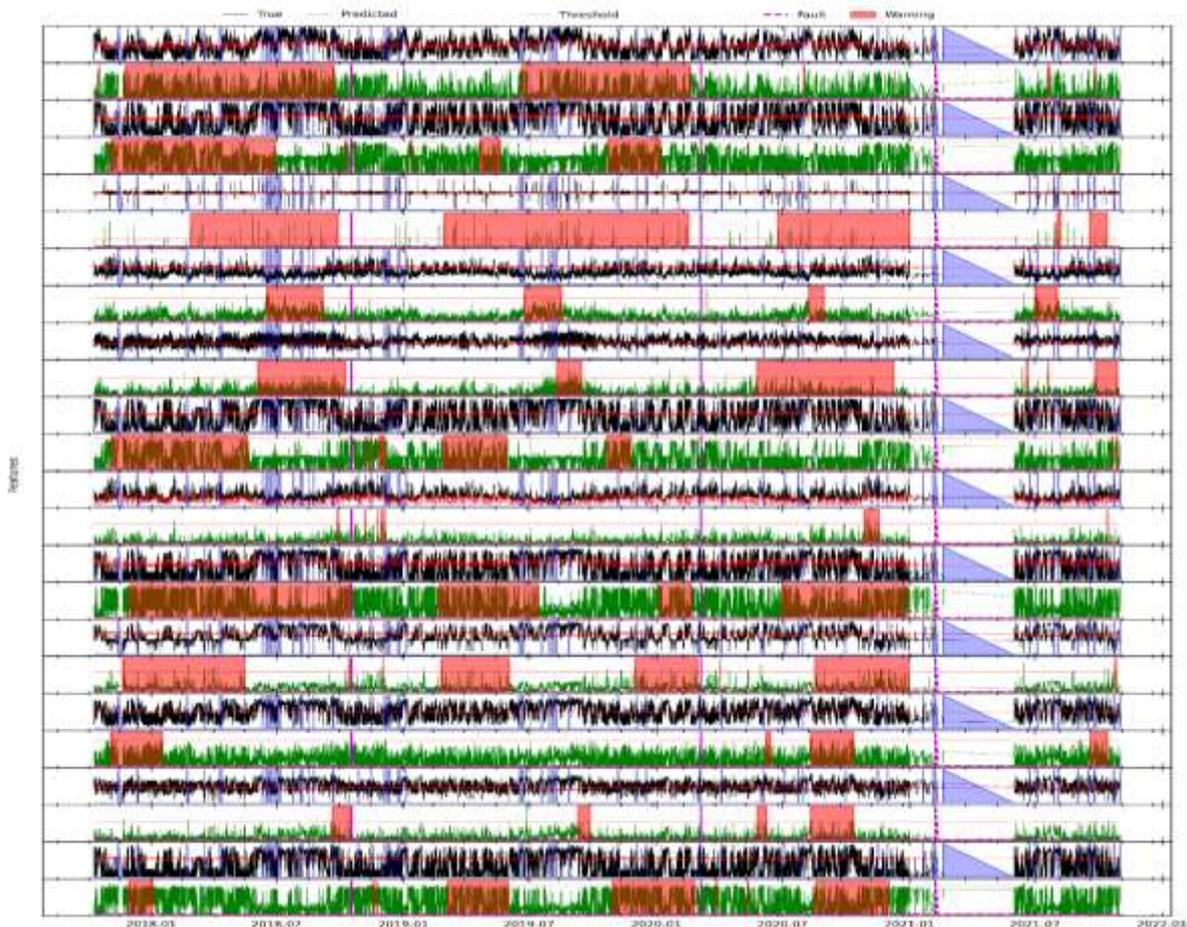


Figura 168. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT7.

En la **Figura 169** en la que se destaca que el modelo advirtió las 3 fallas del módulo IGBT con un avance aproximado de 5 meses de anticipación.

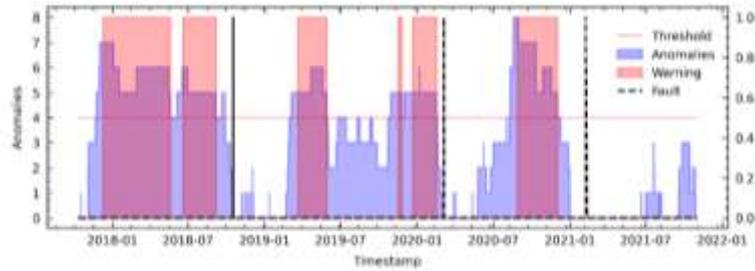


Figura 169. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.

B. TranAD utilizando el método C6 en WT7

En la **Figura 170** se analiza variable a variable la evaluación de fallas de primera etapa.

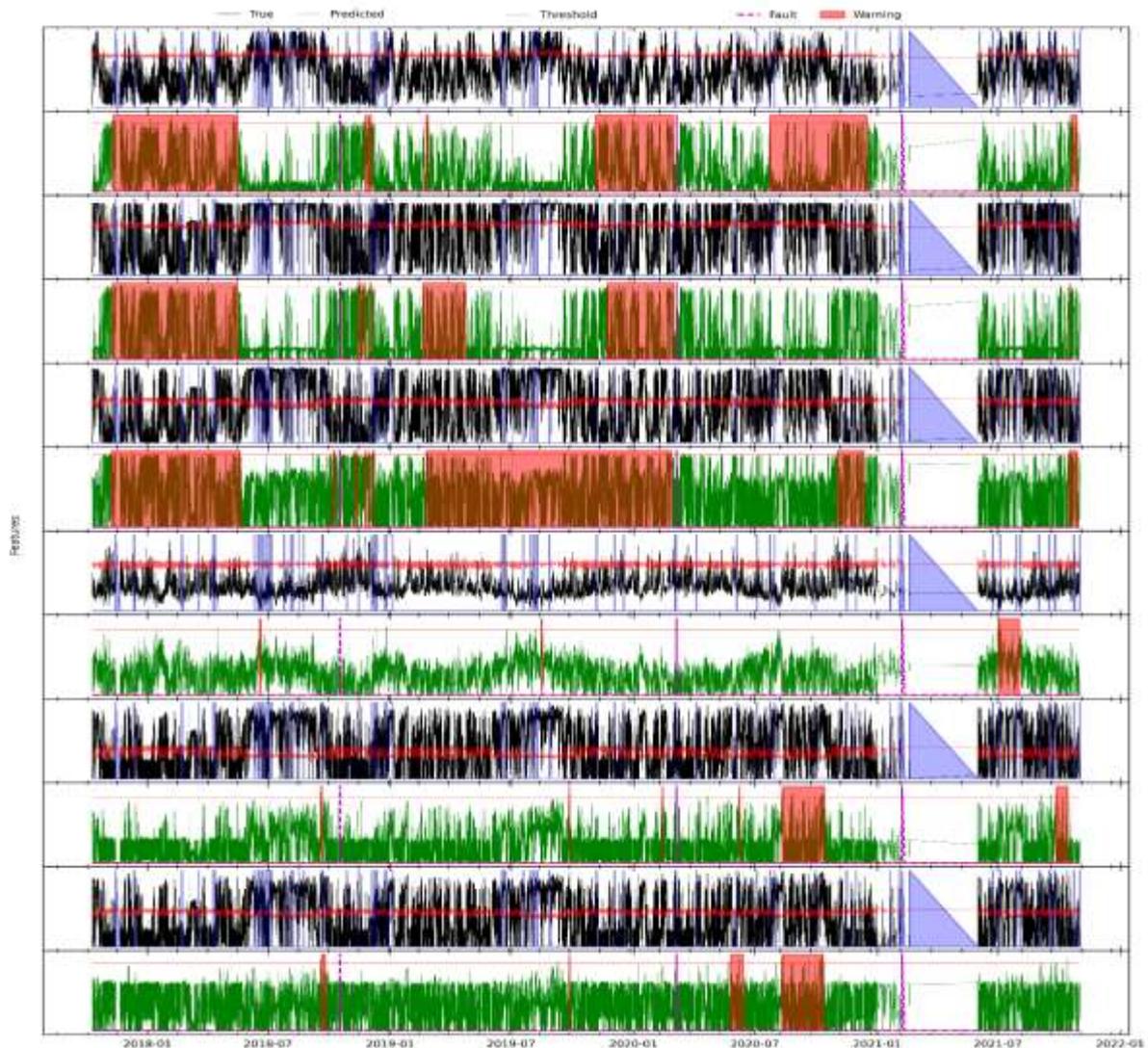


Figura 170. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT7.

En la **Figura 171** se destaca que el modelo advirtió las tres fallas del módulo IGBT con un avance promedio aproximado de 5 meses.

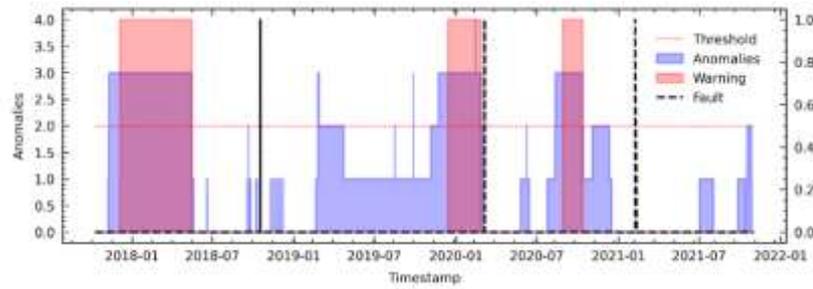


Figura 171. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.

C. TranAD utilizando el método PV18 en WT7

En la **Figura 172** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

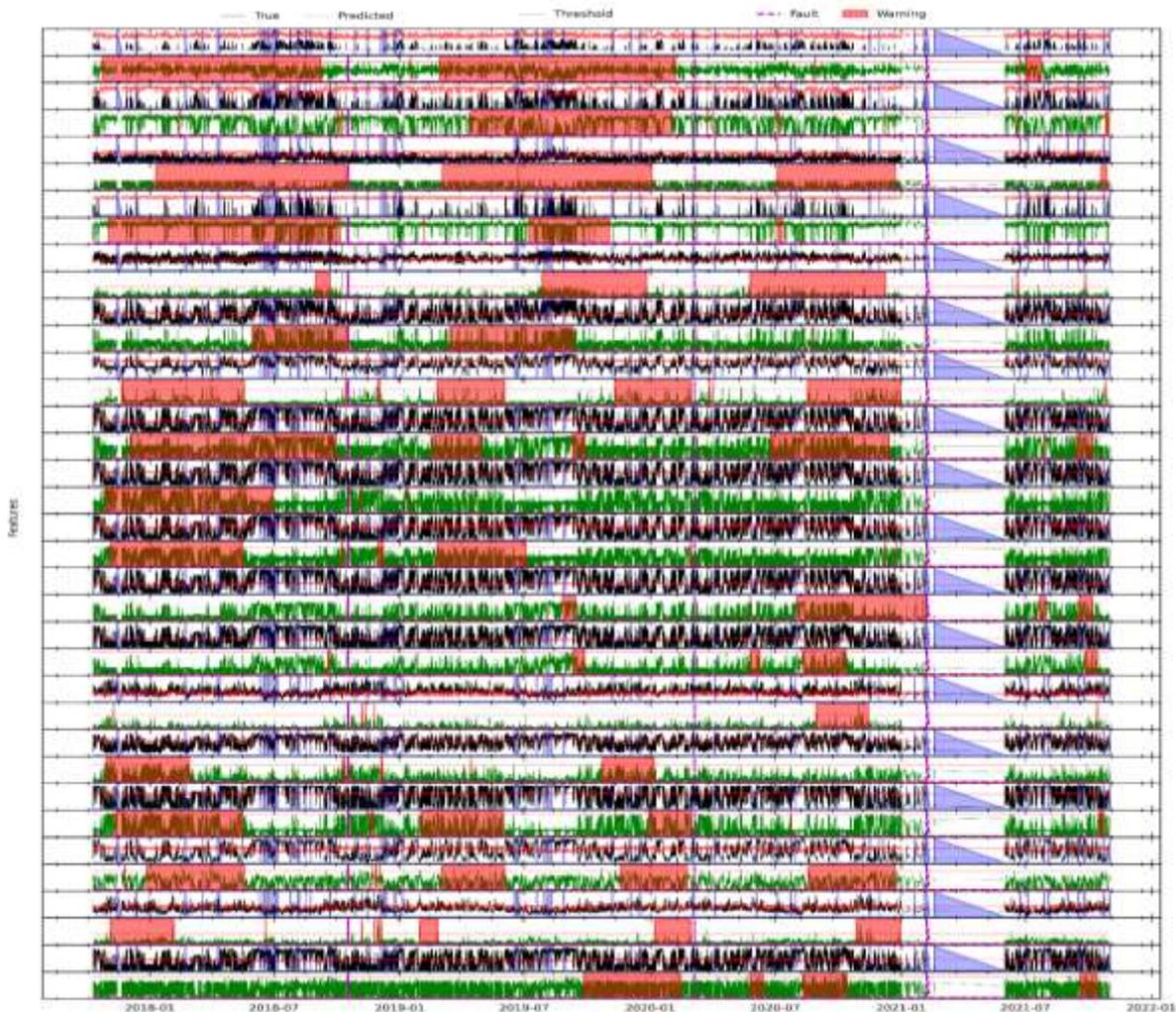


Figura 172. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT7.

En la **Figura 173** se muestra que el modelo logró advertir las tres fallas en el módulo IGBT con un avance aproximado de 6 meses y también registró un FP.

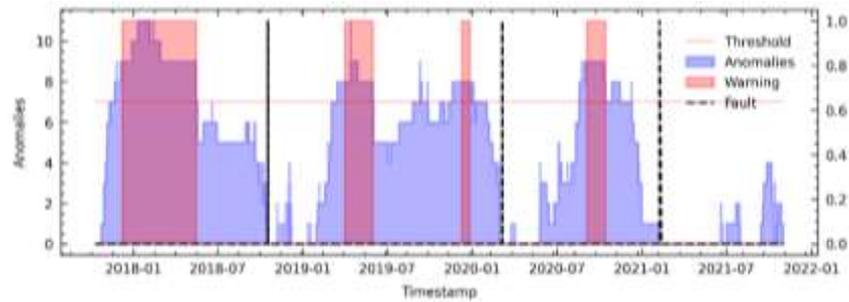


Figura 173. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.

D. TranAD utilizando el método LM18 en WT7

Para el análisis de LM18, en primer lugar, evalúa variable a variable como se observa en la **Figura 174**.

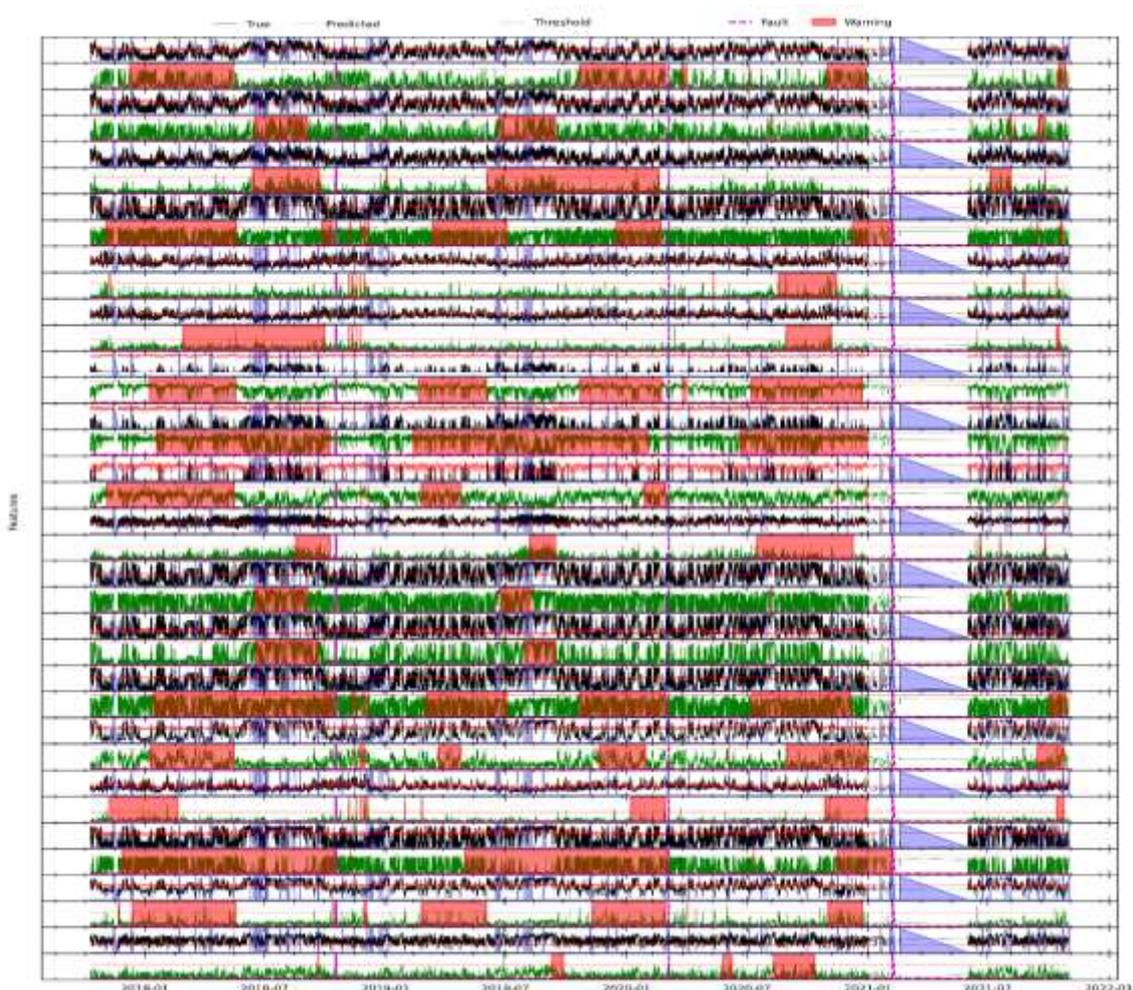


Figura 174. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para la WT7.

En los resultados para LM18 se obtuvo la **Figura 175**, donde se observa que el modelo advirtió las tres fallas con un avance aproximado de 5 meses.

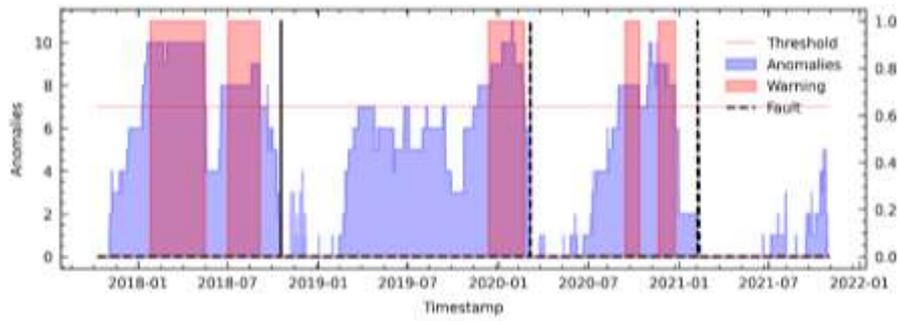


Figura 175. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT7.

XVIII. Detección de fallas en WT8 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT8

En la **Figura 176** se analiza variable a variable la evaluación de fallas de primera etapa.

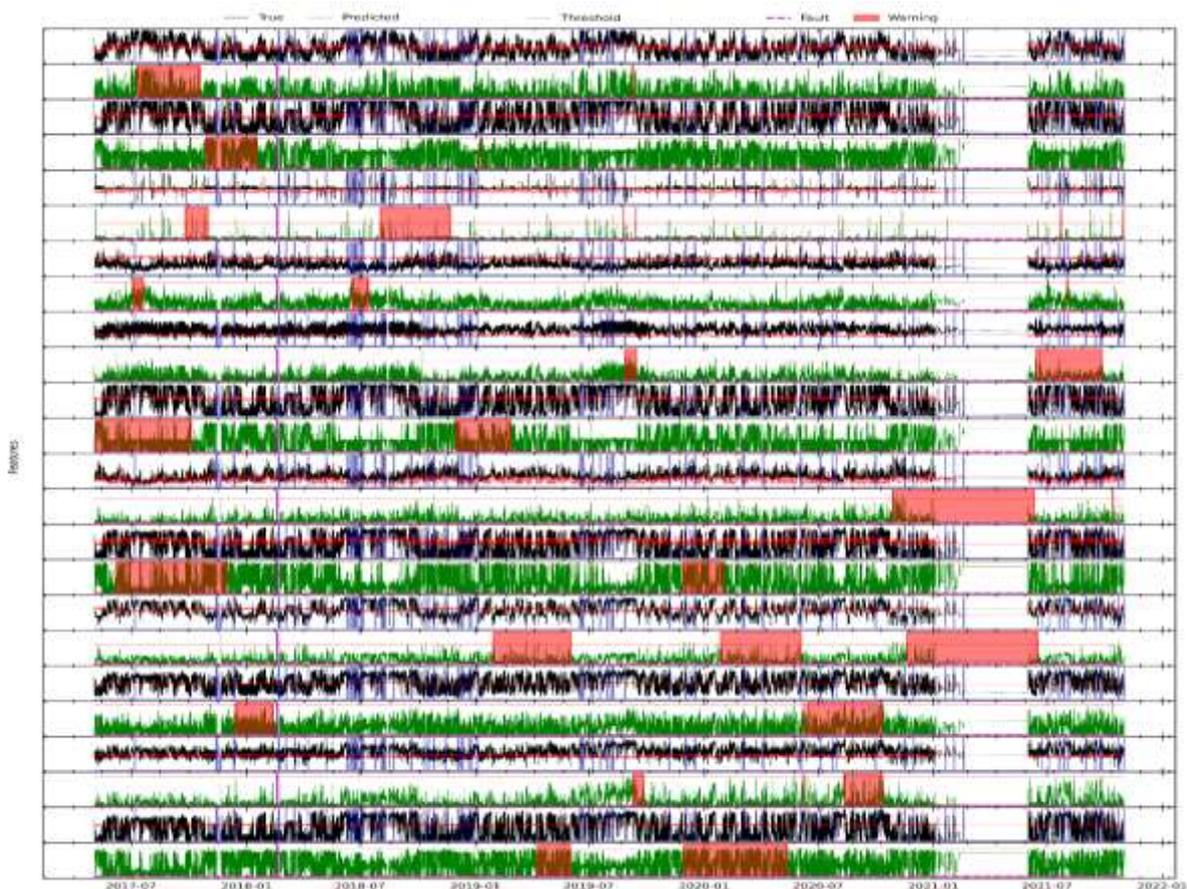


Figura 176. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.

En la **Figura 177** se destaca que el modelo advirtió la falla del módulo IGBT con un avance aproximado de 6 meses antes de su ocurrencia.

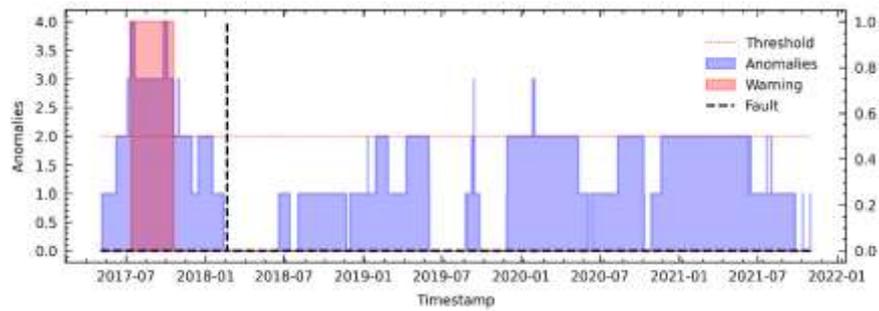


Figura 177. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.

B. TranAD utilizando el método C6 en WT8

En la **Figura 178** se analiza variable a variable la evaluación de fallas de primera etapa.

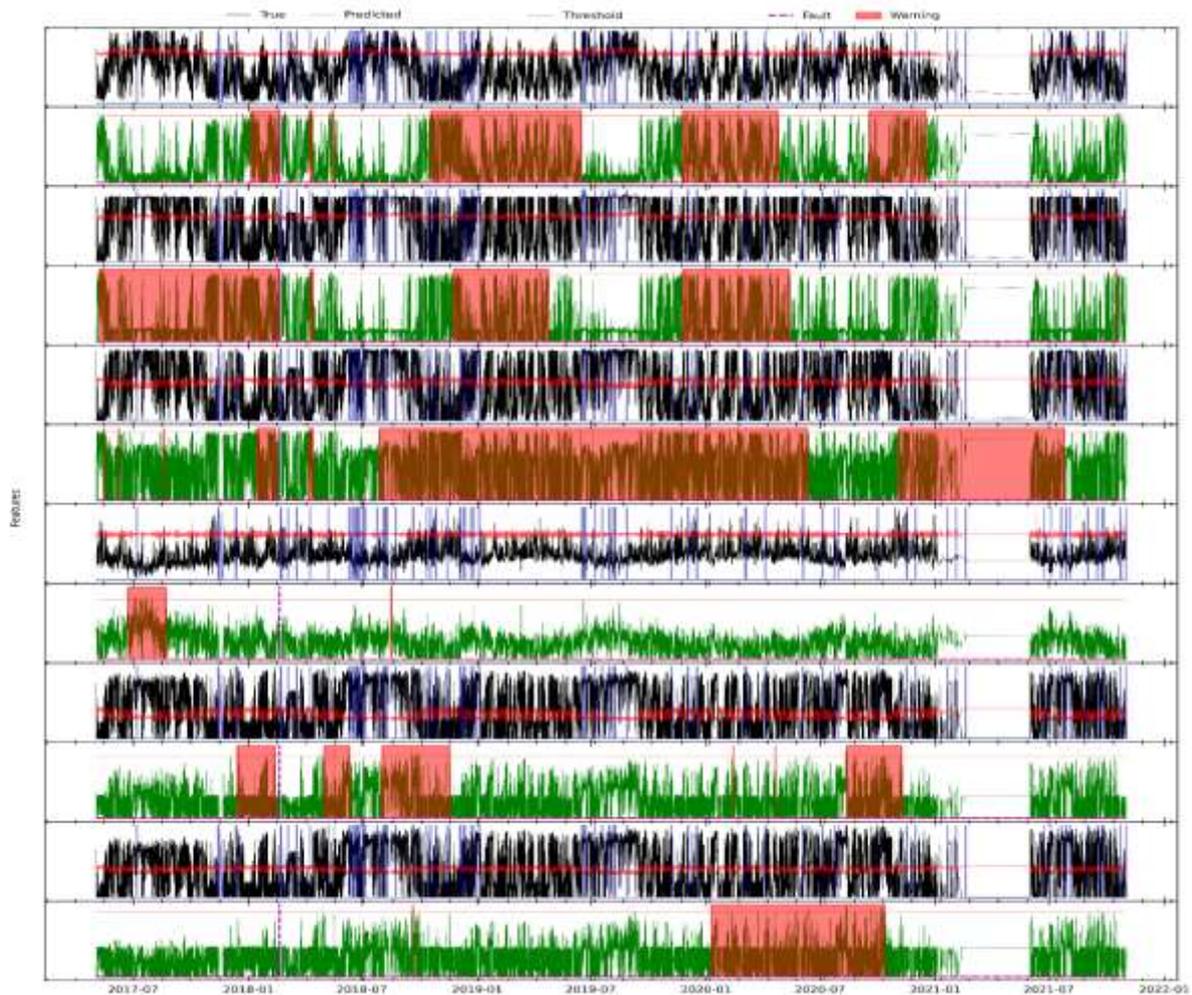


Figura 178. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.

En la **Figura 179** en la que se destaca que el modelo logró advertir la falla con un avance de 1 mes de anticipación y también registró 1 FP a inicios de 2020.

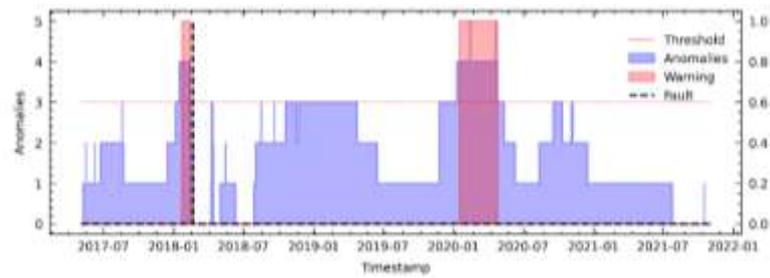


Figura 179. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.

C. TranAD utilizando el método PV18 en WT8

En la **Figura 180** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.



Figura 180. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.

En la **Figura 181** se muestra que el modelo logró advertir la falla del módulo IGBT con un avance aproximado de 2 meses antes de su ocurrencia.

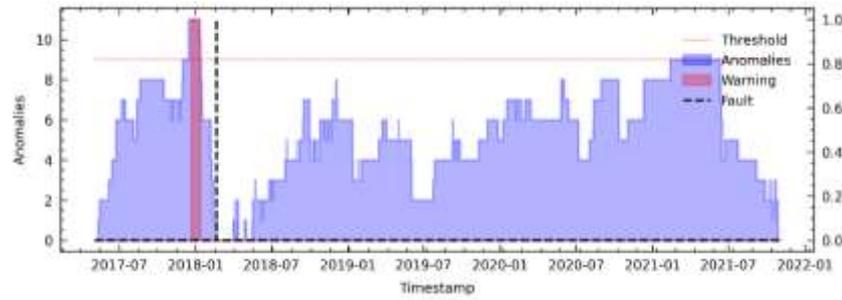


Figura 181. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.

D. TranAD utilizando el método LM18 en WT8

Para el análisis de LM18, en primer lugar, se realiza la evaluación de primera etapa como se muestra en la **Figura 182**.

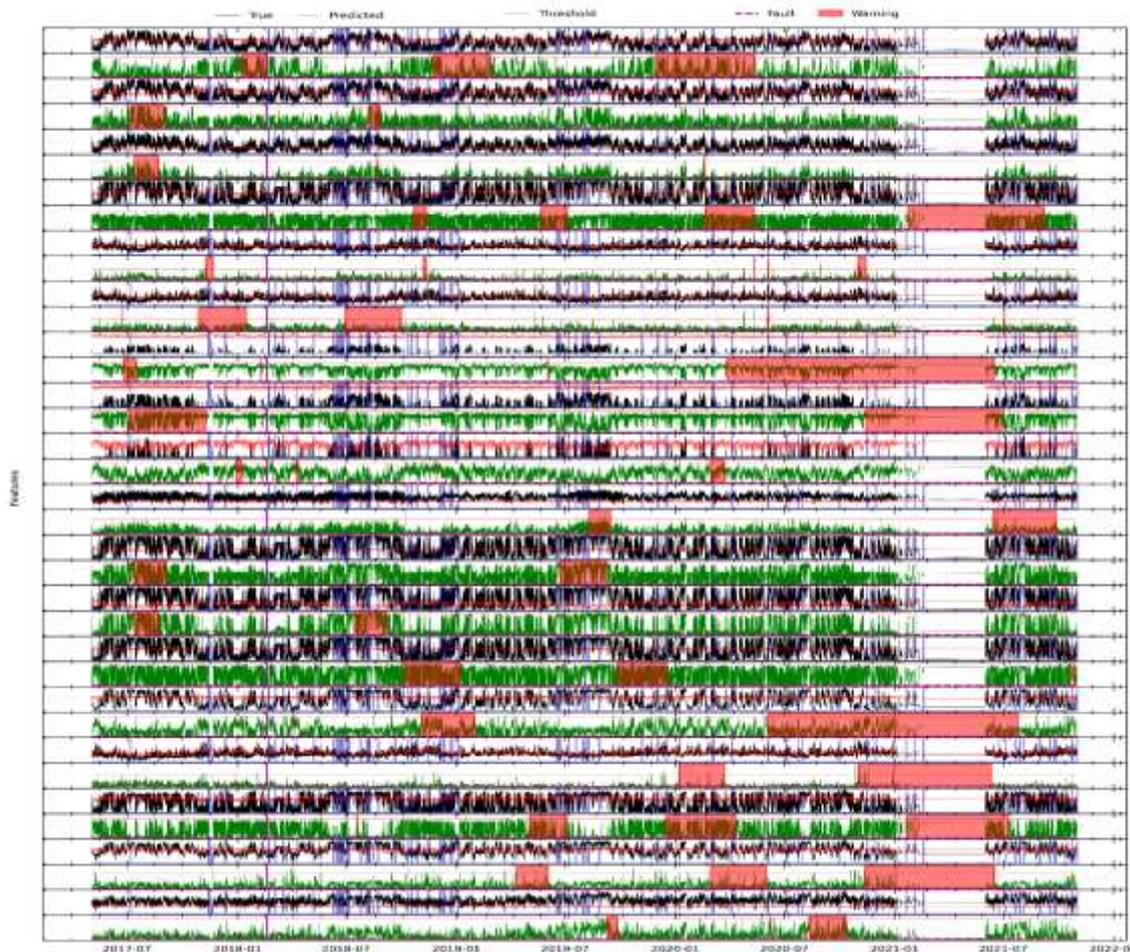


Figura 182. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT8.

En los resultados para LM18 se obtuvo la **Figura 183**, donde se observa que el modelo advirtió la falla con avance aproximado de 6 meses, así mismo advirtió donde no hubo registro de fallo.

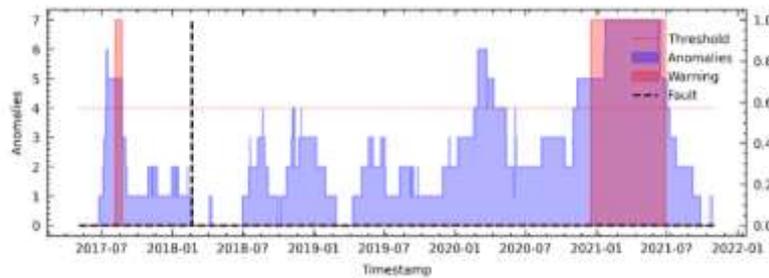


Figura 183. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT8.

XIX. Detección de fallas en WT9 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT9

En la **Figura 184**, se analiza variable a variable la evaluación de fallas de primera etapa.

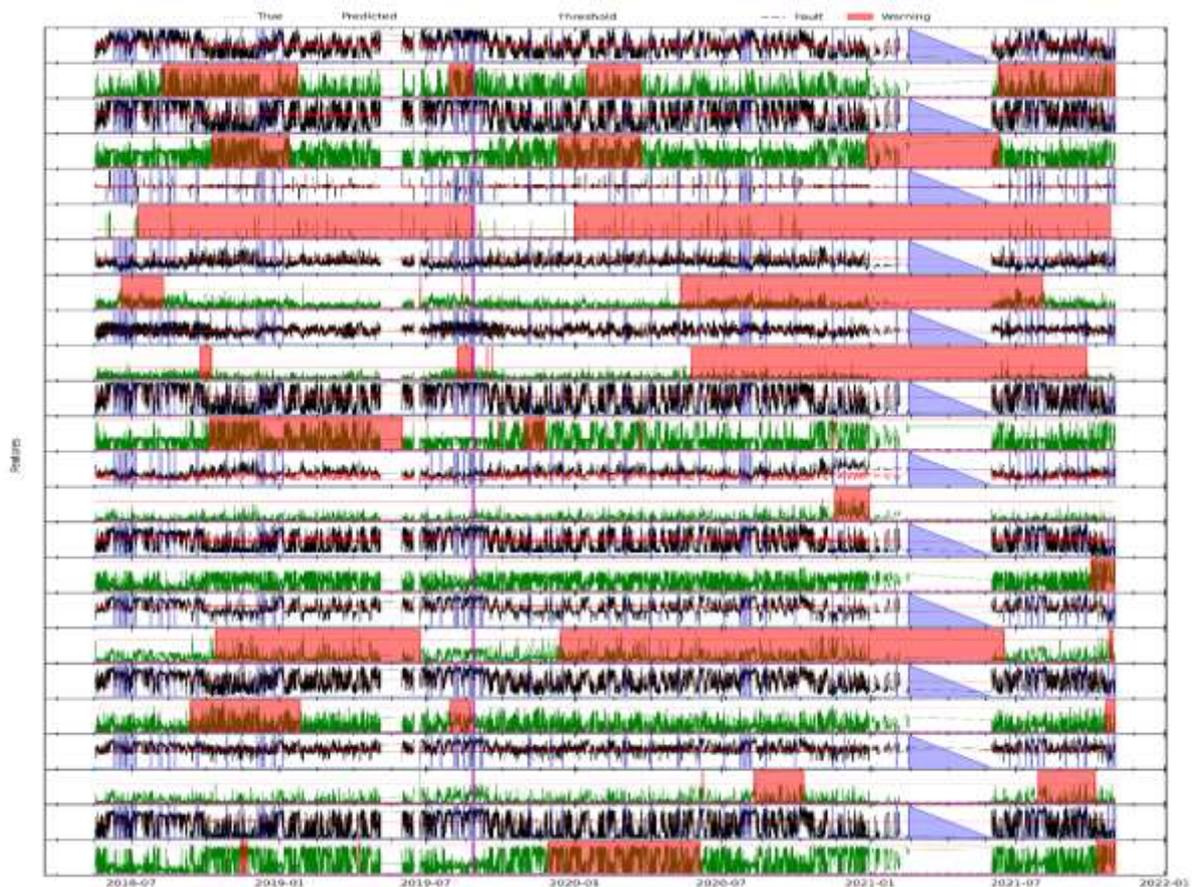


Figura 184. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.

En la **Figura 185** se destaca que el modelo advirtió la falla con un avance aproximado de 8 meses antes de su ocurrencia.

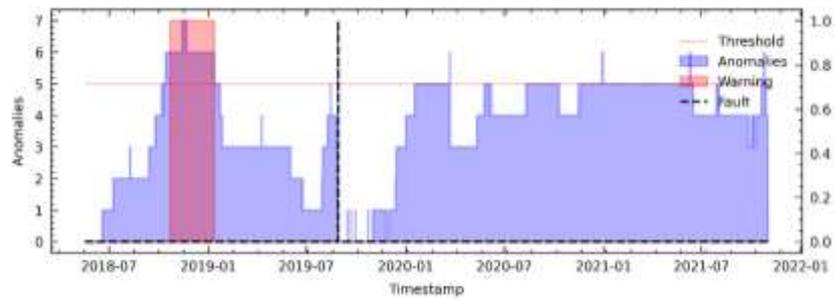


Figura 185. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.

B. TranAD utilizando el método C6 en WT9

En la **Figura 186** se analiza variable a variable la evaluación de fallas de primera etapa.

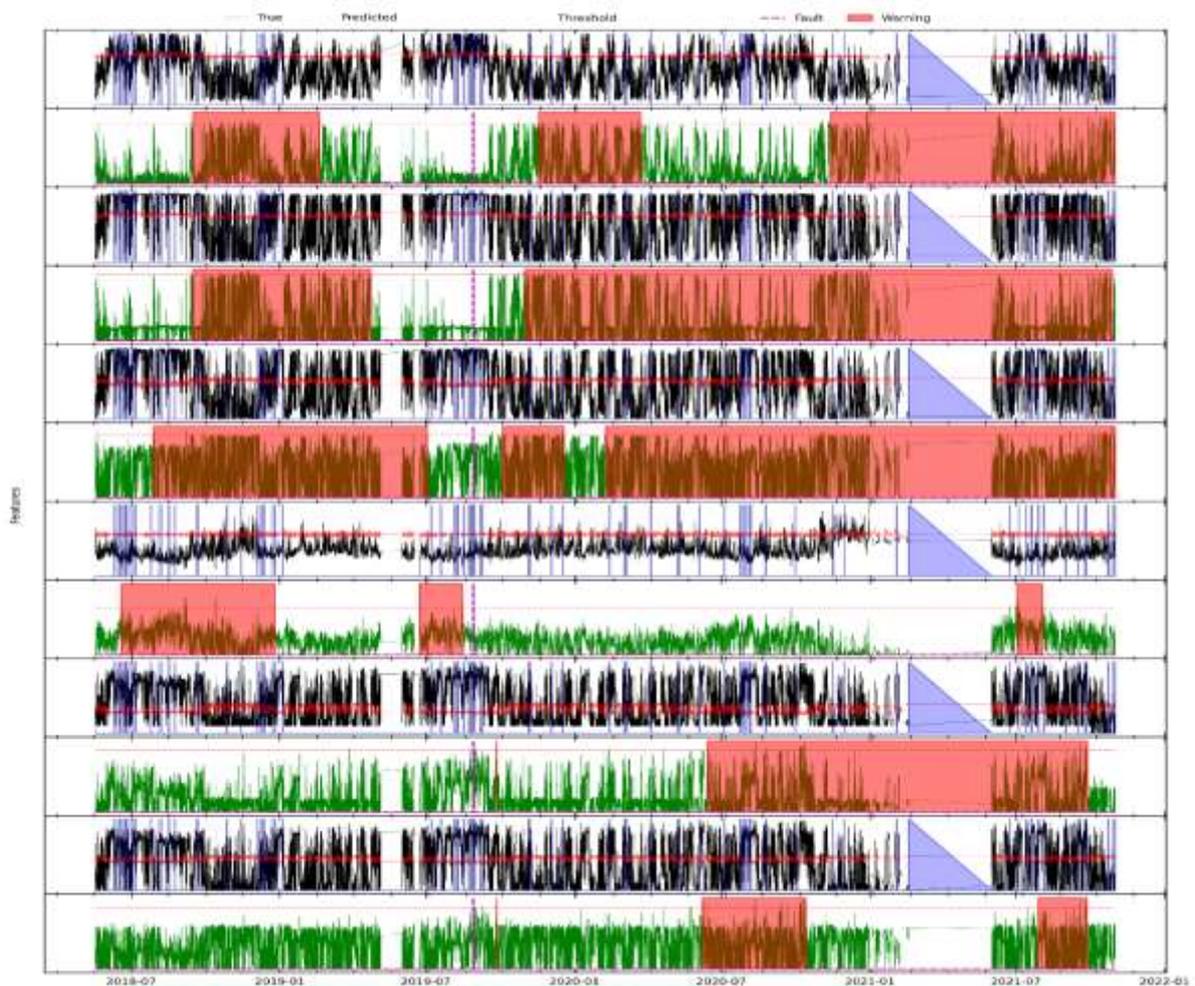


Figura 186. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.

En la **Figura 187** se destaca que el modelo advirtió la falla con un avance aproximado de 8 meses antes de que ocurriese y también registra dos FPs.

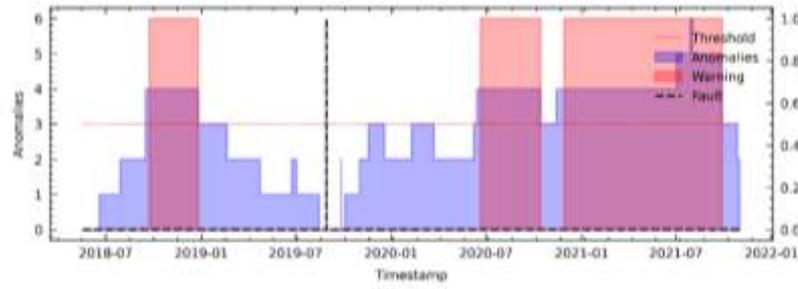


Figura 187. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.

C. TranAD utilizando el método PV18 en WT9

En la **Figura 188** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

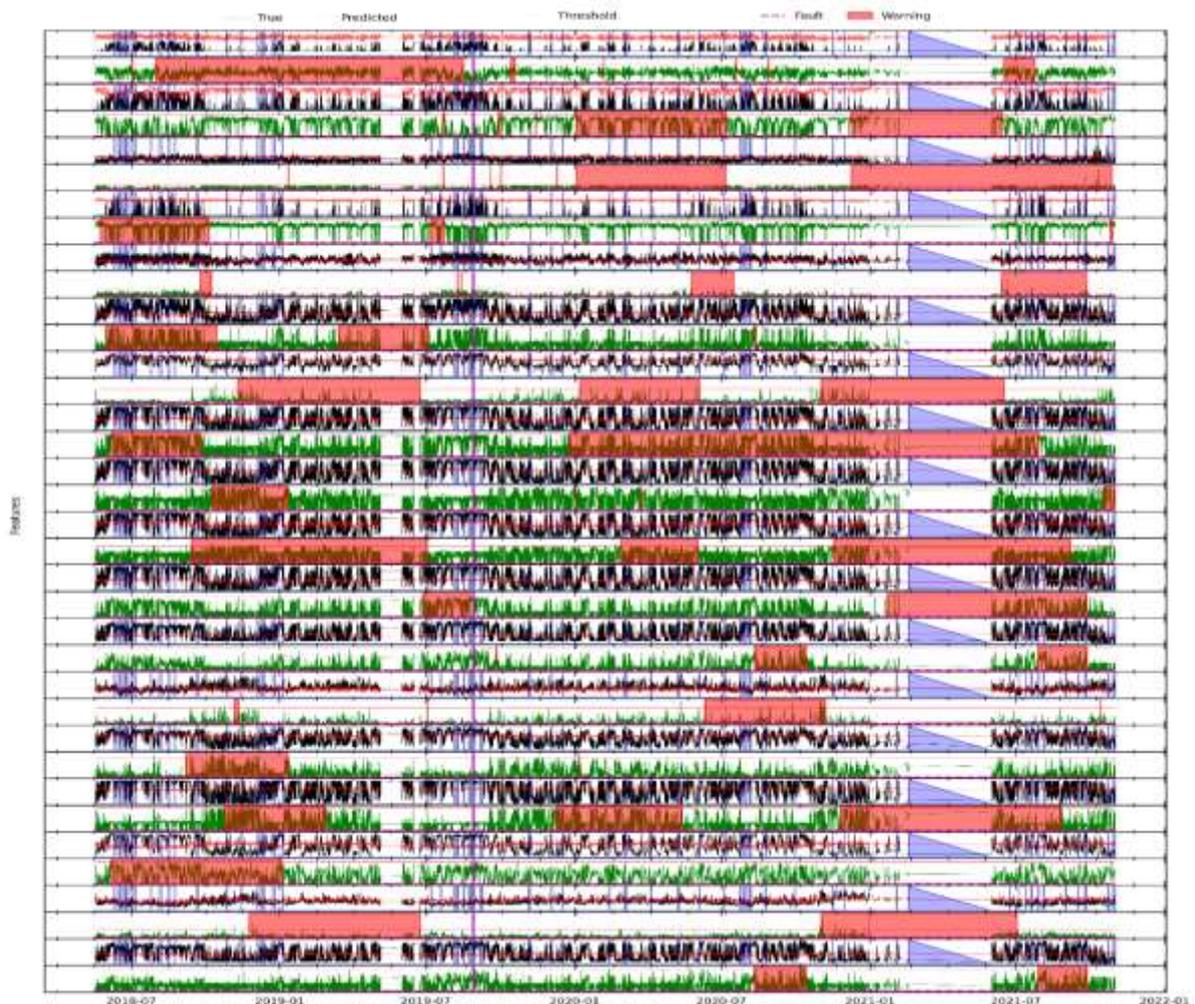


Figura 188. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.

En la **Figura 189** se muestra que el modelo logró advertir la falla en el módulo IGBT con un avance aproximado de 8 meses, pero también registró un FP.

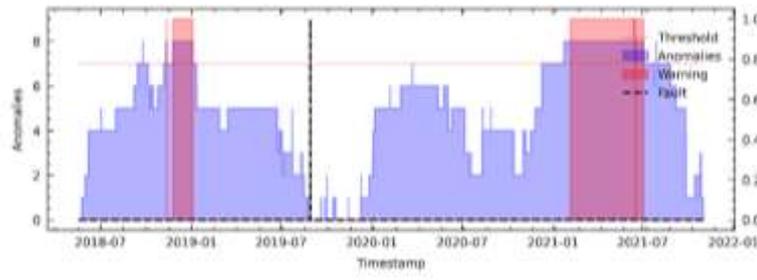


Figura 189. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.

D. TranAD utilizando el método LM18 en WT9

Para el análisis de LM18, se realiza la evaluación de primera etapa como se muestra en la **Figura 190**.

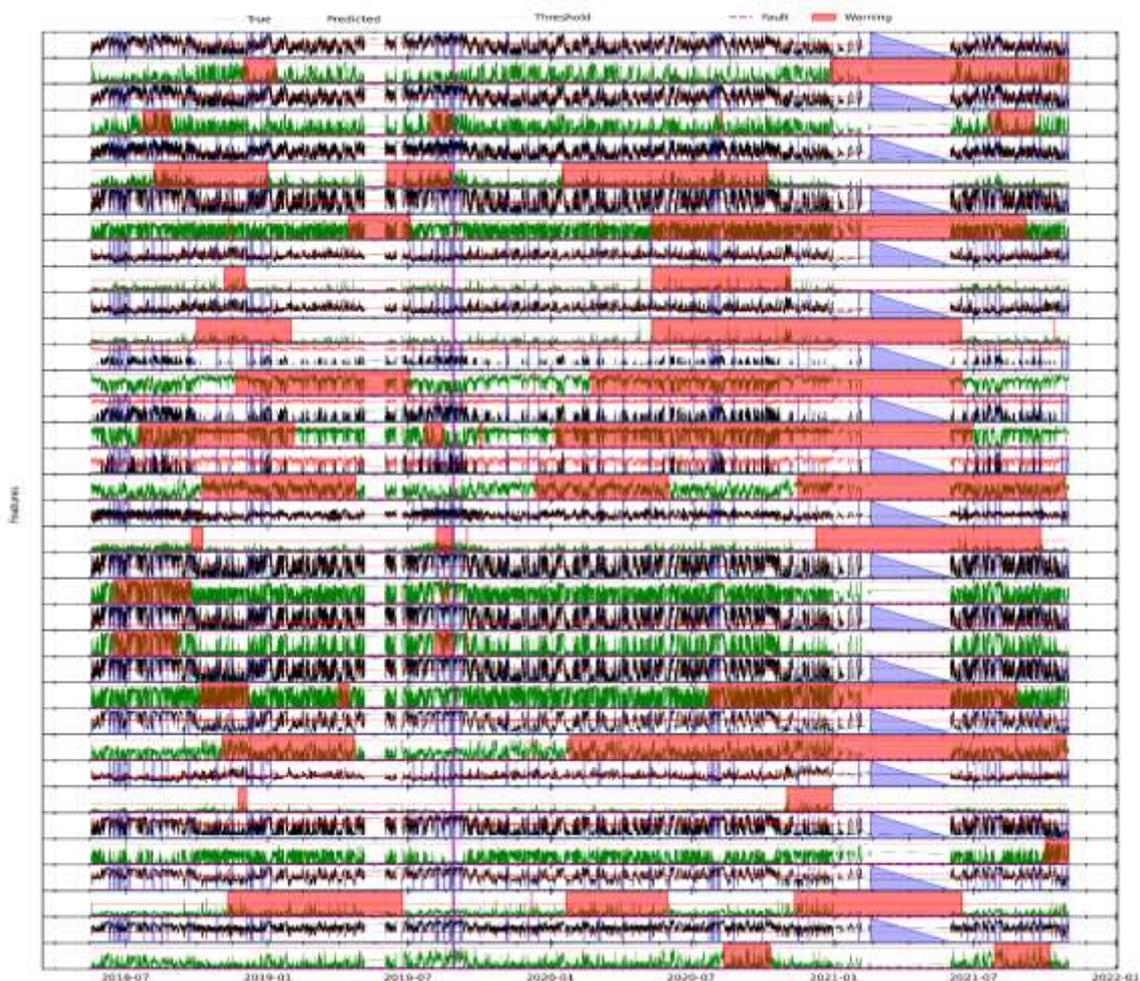


Figura 190. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT9.

En los resultados para LM18 se obtuvo la **Figura 191**, donde se observa que el modelo advirtió la falla con un avance aproximado de 8 meses antes de su ocurrencia, así mismo advirtió en tres ocasiones donde no ocurrió falla.

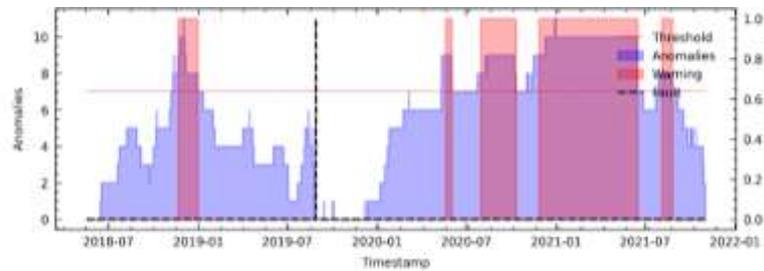


Figura 191. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT9.

XX. Detección de fallas en WT10 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT10

En la **Figura 192** se analiza la evaluación de fallas de primera etapa.

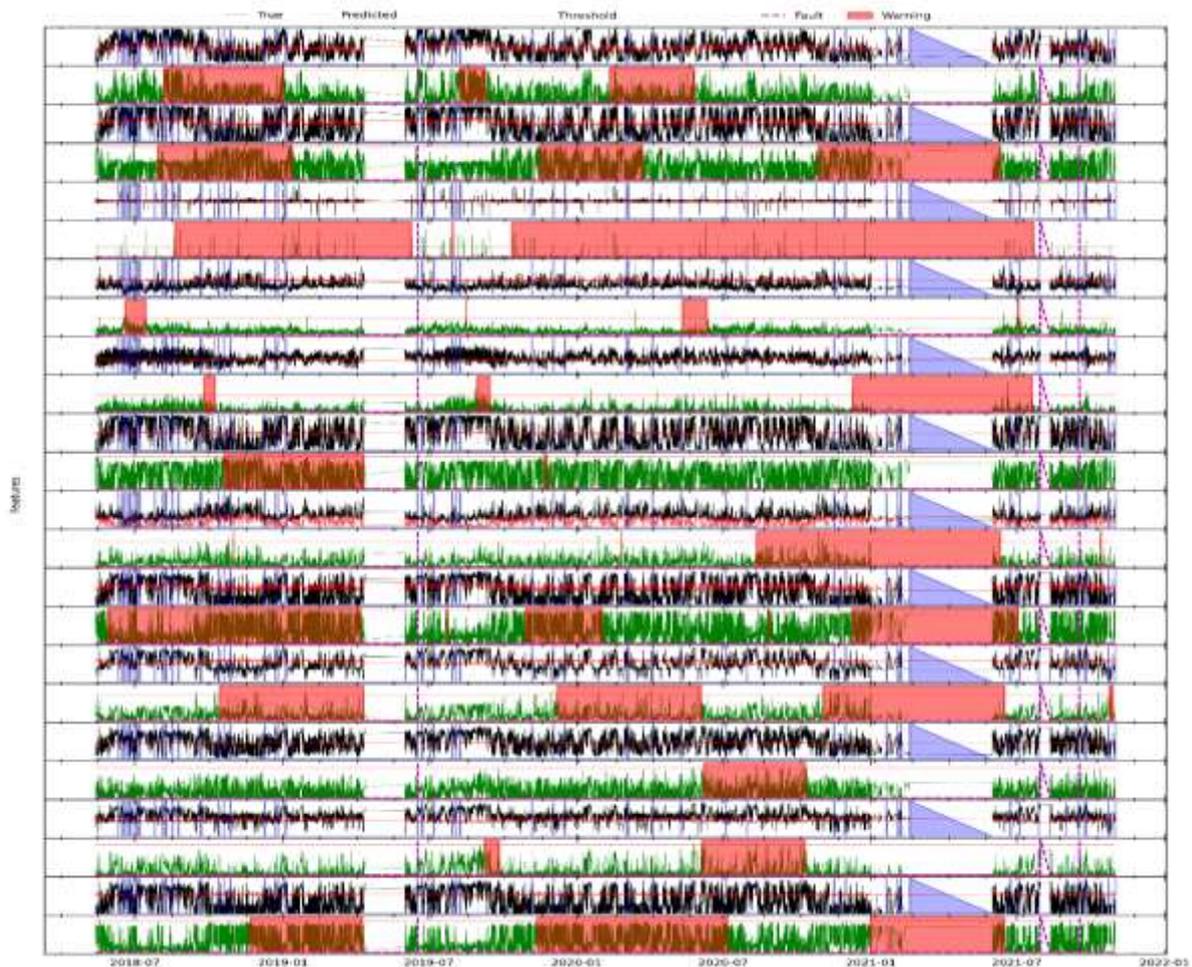


Figura 192. Resultados para RFE12 en el análisis de advertencias en el módulo IGBT mediante el modelo TranAD para WT10.

En la **Figura 193** se destaca que el modelo advirtió las dos fallas del módulo IGBT con un avance aproximado de 7 meses antes de que ocurra la falla.

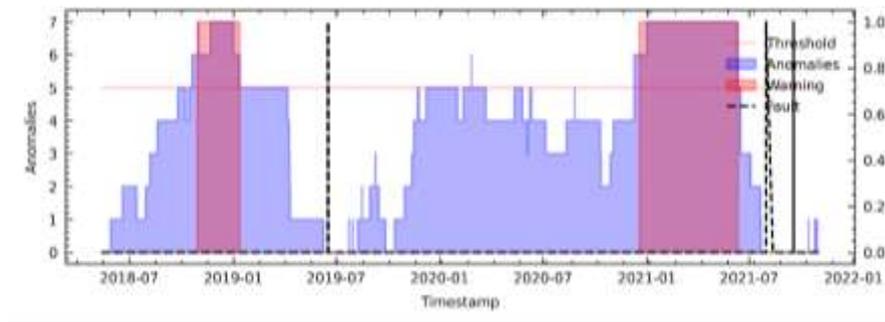


Figura 193. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.

B. TranAD utilizando el método C6 en WT10

En la **Figura 194** se analiza variable a variable la evaluación de fallas de primera etapa.

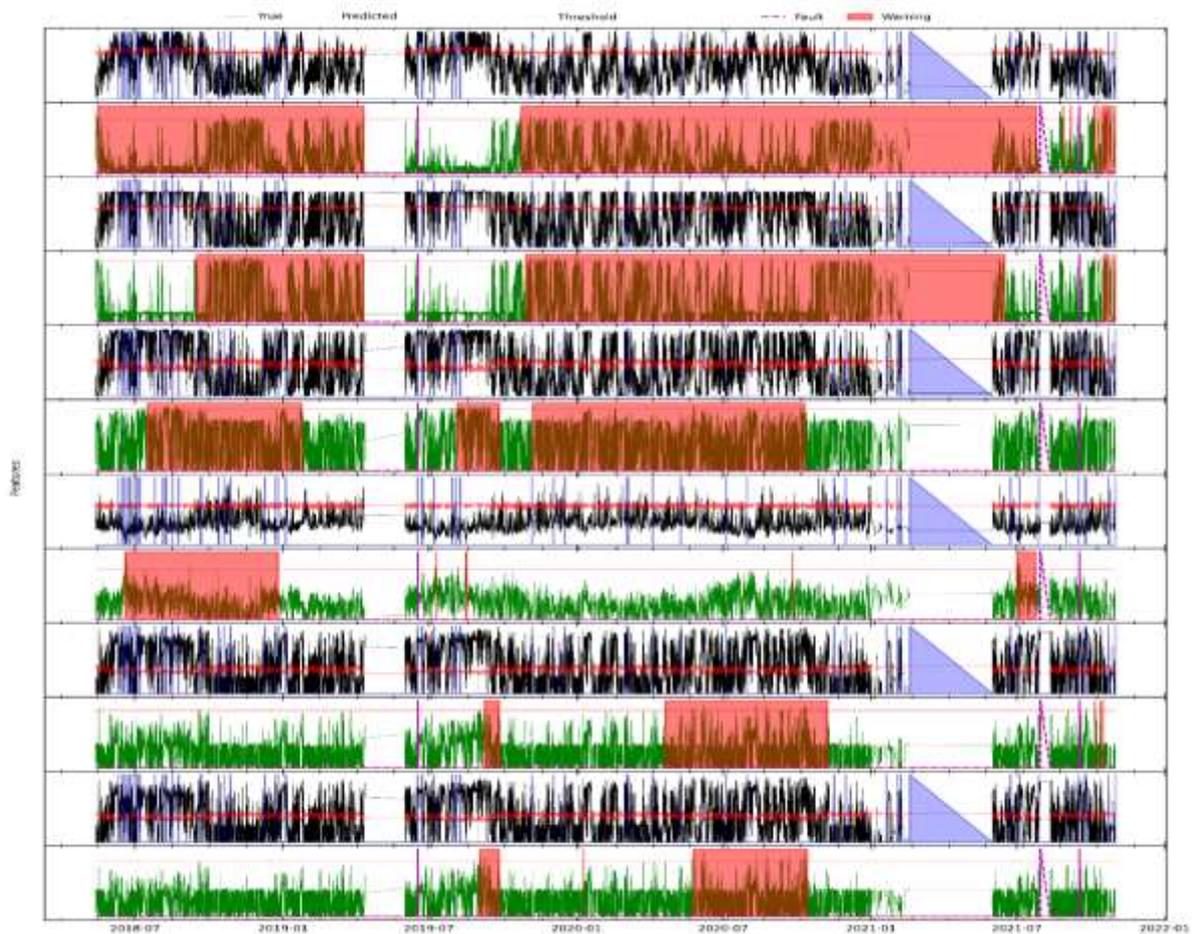


Figura 194. Resultados para C6 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT10.

En la **Figura 195** se destaca que el modelo advirtió una de las dos fallas en el módulo IGBT con un avance aproximado de 8 meses y registra 1 FP.

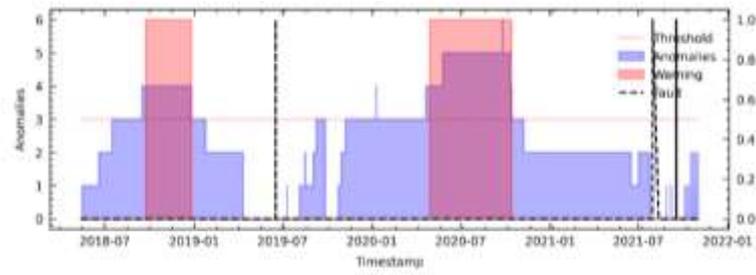


Figura 195. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.

C. TranAD utilizando el método PV18 en WT10

En la **Figura 196** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.



Figura 196. Resultados para PV18 en el análisis advertencias variable a variable mediante el modelo TranAD para WT10.

En la **Figura 197**, se muestra que el modelo logró advertir las dos fallas en el módulo IGBT con un avance aproximado 7 meses, pero también registró un FP.

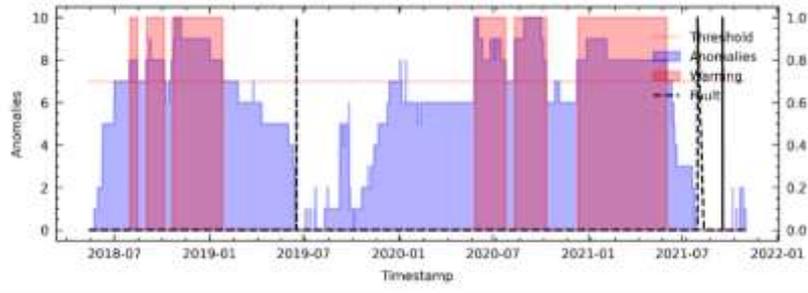


Figura 197. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.

D. TranAD utilizando el método LM18 en WT10

Para el análisis de LM18, en primer lugar, se debe evaluar las fallas de primera etapa como se muestra en la **Figura 198**.

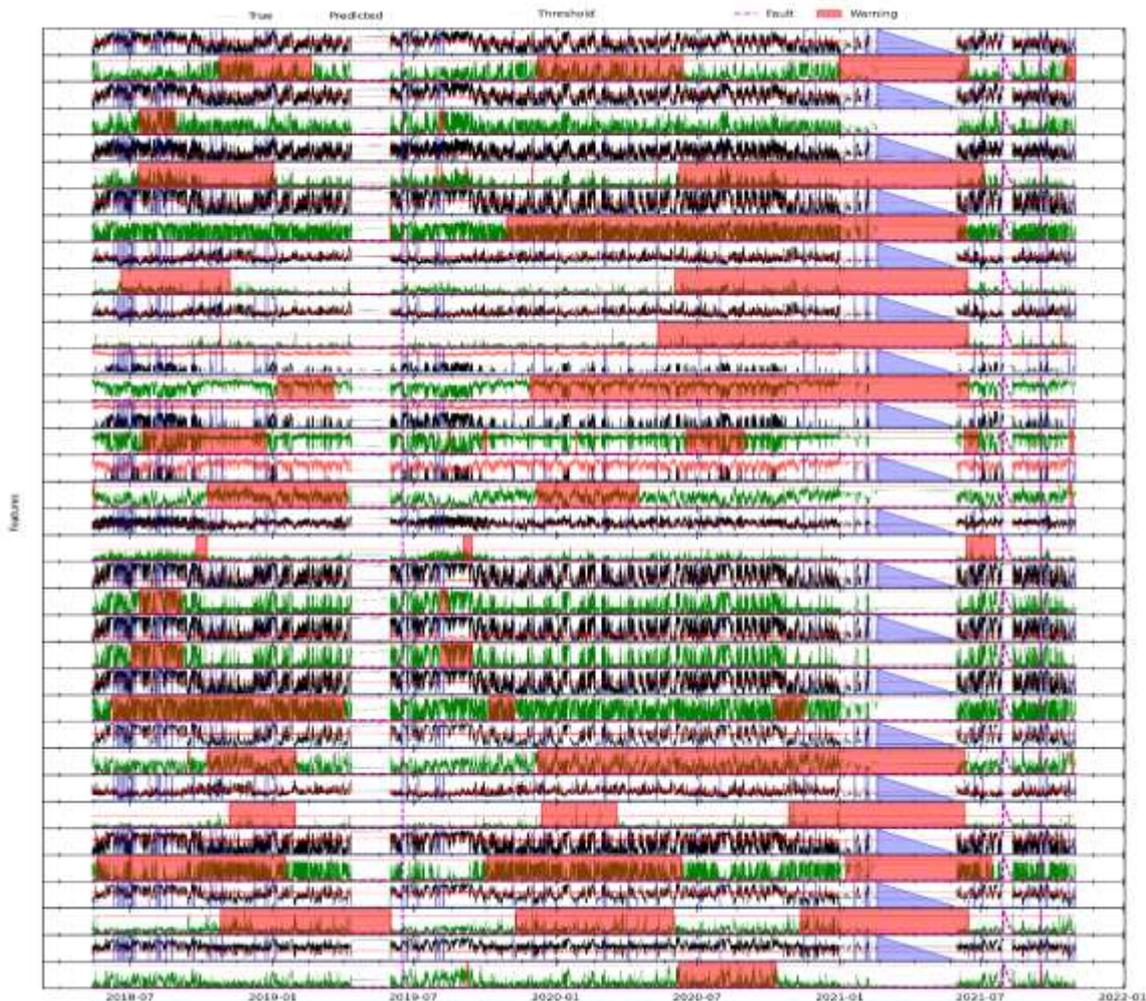


Figura 198. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT10.

En los resultados para LM18 se obtuvo la **Figura 199**, donde se observa que el modelo advirtió las dos fallas con un avance aproximado de 6 meses antes de su ocurrencia, así mismo advirtió en una ocasión donde no ocurrió fallo.

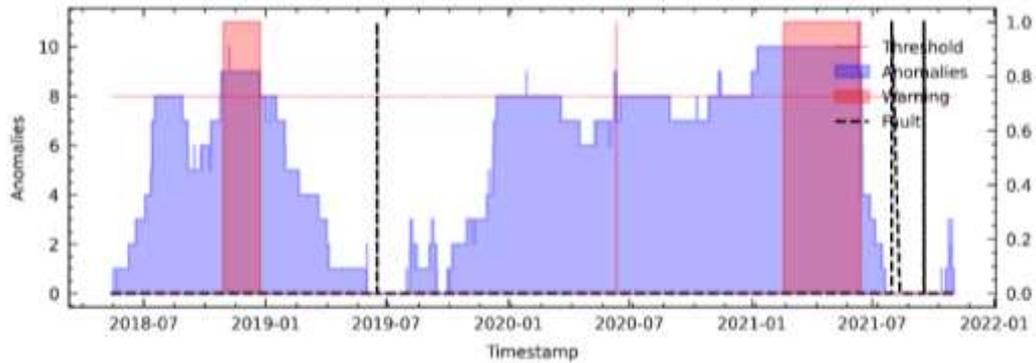


Figura 199. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT10.

XXI. Detección de fallas en WT11 mediante TranAD (detalladamente).

A. TranAD utilizando el método RFE12 en WT11

En la **Figura 200** se analiza variable a variable la evaluación de falla de primera etapa.

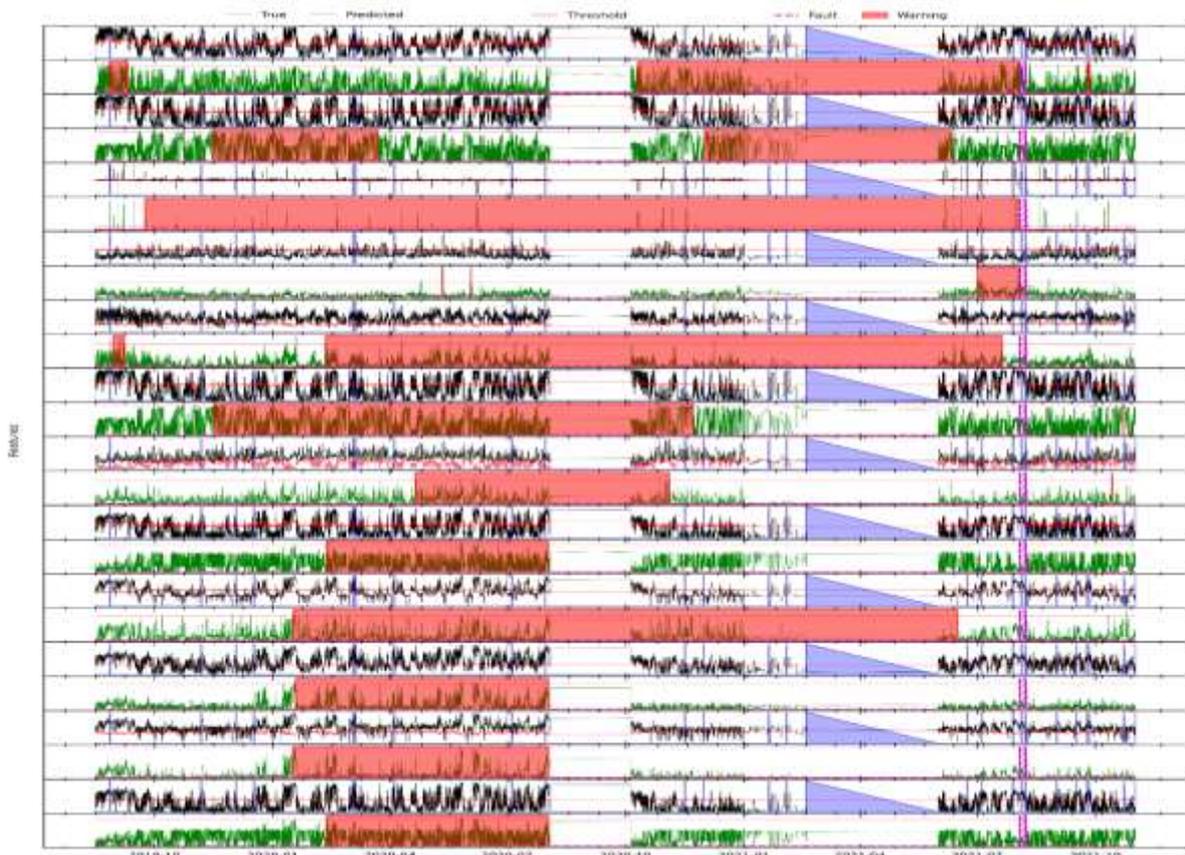


Figura 200. Resultados para RFE12 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT11.

En la **Figura 201**, se destaca que el modelo no logró advertir la primera falla registrada en el archivo de O&M.

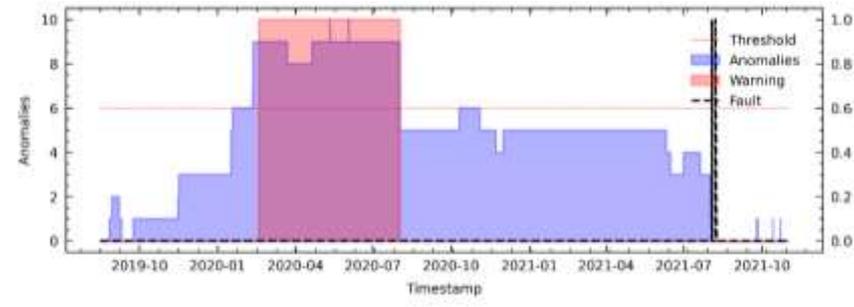


Figura 201. Resultados para RFE12 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.

B. TranAD utilizando el método C6 en WT11

En la **Figura 202** se analiza variable a variable la evaluación de fallas de primera etapa.

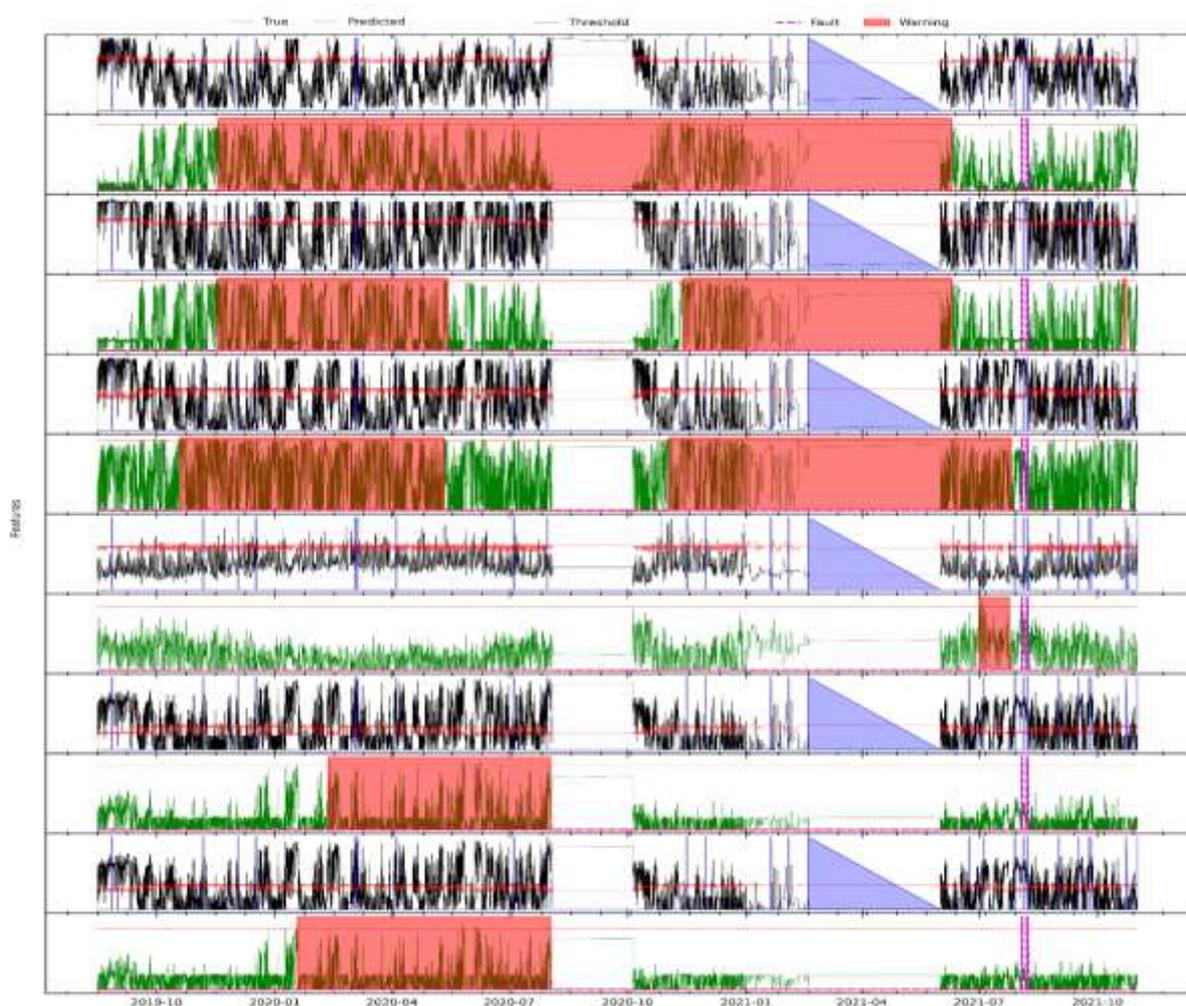


Figura 202. Resultados para C6 en el análisis de advertencias en el módulo IGBT mediante el modelo TranAD para WT11.

En la **Figura 203** se destaca que el modelo no logró advertir la falla del módulo IGBT y registró un FP.

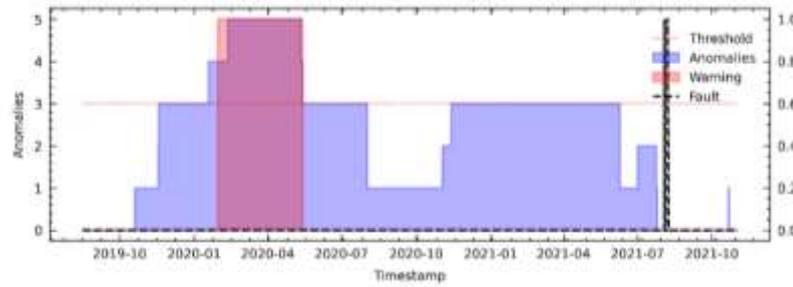


Figura 203. Resultados para C6 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.

C. TranAD utilizando el método PV18 en WT11

En la **Figura 204** se analiza variable por variable las advertencias de falla para su posterior análisis en la segunda etapa.

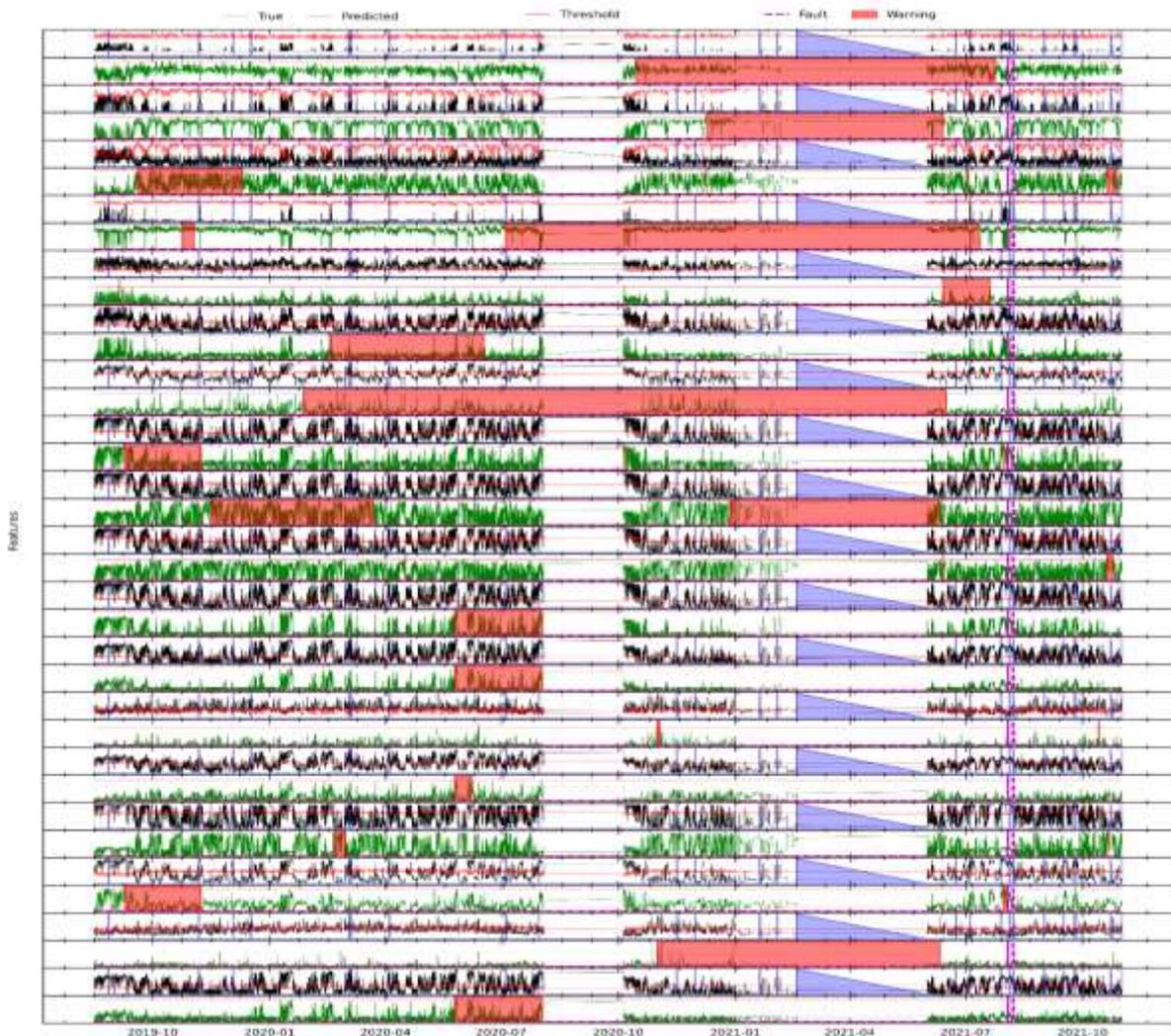


Figura 204. Resultados para PV18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT11.

En la **Figura 205**, se muestra que el modelo logró advertir la falla del módulo IGBT con un avance aproximado 5 meses de anticipación, pero también logró un FN.

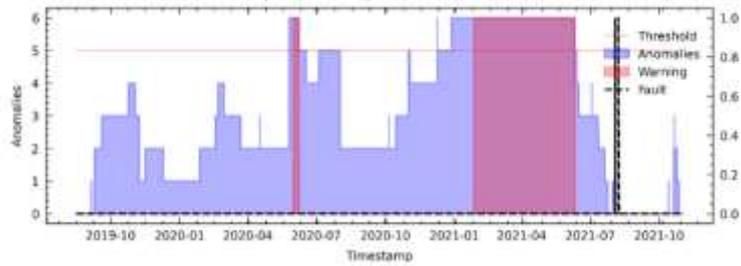


Figura 205. Resultados para PV18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.

D. TranAD utilizando el método LM18 en WT11

Para el análisis de LM18, en primer lugar, se realizó la evaluación de primera etapa como se muestra en la **Figura 206**.

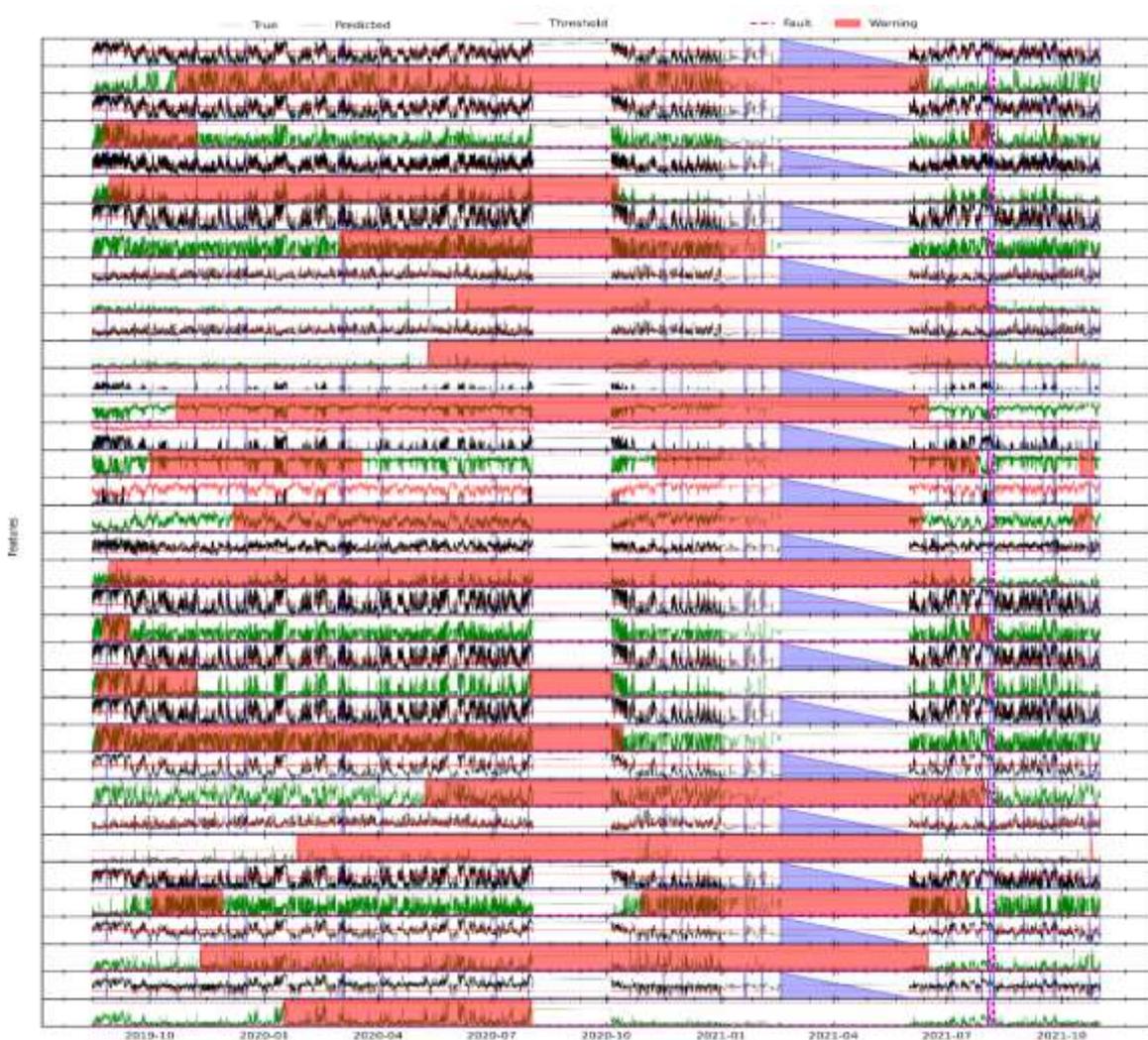


Figura 206. Resultados para LM18 en el análisis de advertencias variable a variable mediante el modelo TranAD para WT11.

En los resultados para LM18 se obtuvo la **Figura 207**, donde se observa que el modelo advirtió la falla del módulo IGBT con un avance aproximado 6 meses antes de su ocurrencia, así mismo registró un FP.

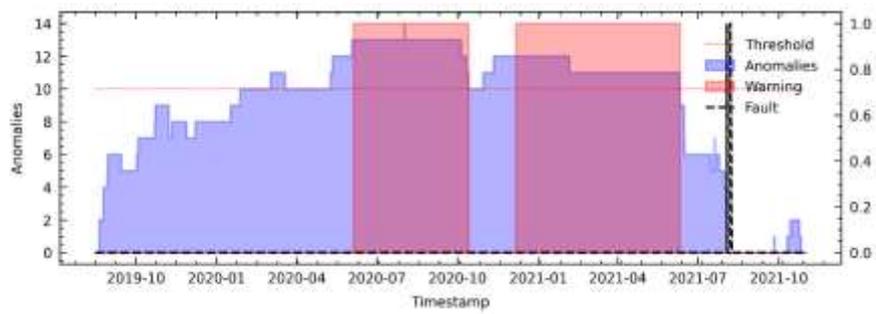


Figura 207. Resultados para LM18 en el análisis de fallas del módulo IGBT mediante el modelo TranAD para WT11.

Anexo 5. Guía para ejecutar de manera efectiva los resultados de este trabajo

Después de solicitar el código fuente a los responsables descritos en la sección de recomendaciones, recibirás una carpeta titulada 'Python Torres Cabrera Joel Stiven Torres'. Deberás cargar esta carpeta a tu Google Drive e instalar Google Colab. Es importante destacar que algunos procesos pueden requerir un mínimo de 30 GB de memoria RAM, para lo cual puedes utilizar Google Colab Pro.

Además, ten en cuenta que esta guía puede estar sujeta a cambios. En la carpeta compartida encontrarás un archivo de texto con la guía actualizada, así como posibles nuevas versiones.

Para el procesamiento, etiquetado y filtrado de los datos se deben seguir los siguientes pasos:

1. Para cargar los datos, dirígete a la carpeta 'Python Torres Cabrera Joel Stiven Torres/Dataset original/archivos m'. En esta ubicación deberás cargar los archivos obtenidos.

 2014	 U1	 m140101.txt 
 2015	 U2	 m140102.txt 
 2016	 U3	 m140103.txt 
 2017	 U4	 m140104.txt 
 2018	 U5	 m140105.txt 
 2019	 U6	 m140106.txt 
 2020	 U7	 m140107.txt 
 2021	 U8	 m140108.txt 
	 U9	 m140109.txt 
	 U10	 m140110.txt 
	 U11	 m140111.txt 
		 m140112.txt 

Figura 208. Archivos de dataset de operación de los aerogeneradores.

2. Además, carga los archivos de alarmas del SCADA y el archivo de O&M en la ruta 'Python Torres Cabrera Joel Stiven Torres/Dataset original/fallas'. Aquí se muestran los archivos correspondientes.

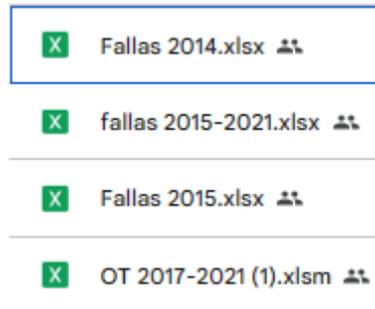


Figura 209. Archivos de alarmas del SCADA y de mantenimiento de los aerogeneradores.

3. Después abrir el archivo "PROCESAMIENTO_DE_DATOS_COLAB_25GB RAM.ipynb" en Google Colab, asegúrate de cambiar la ruta de acceso para que apunte a la carpeta "Python Torres Cabrera Joel Stiven Torres". Esto garantizará que puedas acceder a los archivos necesarios para tu trabajo.

```
[ ] 1 #Aquí debe cambiar su ruta de archivo, por la ubicación de la carpeta "Python Torres Cabrera Joel Stiven Torres" descargada.  
2 Drive-"/content/drive/MyDrive/Python Torres Cabrera Joel Stiven Torres"
```

Figura 210. Celda de código de ruta de archivos para procesamiento de datos.

4. Después de ejecutar todas las celdas, obtendrás el archivo "dataset_completo_V2_etiquetado.csv". Este archivo estará disponible en la carpeta "Python Torres Cabrera Joel Stiven Torres/Dataset original". Contendrá todos los datos extraídos y estará listo para su análisis posterior.

- Para la selección de características y visualización de variables y división de los datos.

5. Para proceder, abre el archivo "FSINGIENERIA_CHARACTERISTICAS_COLAB_25GB RAM.ipynb" y asegúrate de cambiar la ruta de acceso de la carpeta de la misma manera que se hizo en el paso 3 anterior. Esto garantizará que puedas trabajar con los datos de manera efectiva en Google Colab. Una vez ejecutado el cuaderno se descargarán datasets de entrenamiento.

- Entrenamiento y prueba del modelo Anomaly Transformer

6. Para entrenar el Modelo Anomaly Transformer modificado en este estudio, primero, acceda a la carpeta del modelo llamada "Anomaly-Transformer-modified_Torres-20240418T150717Z-001". Luego, abra el archivo "Ejecución_anomali_transformer.ipynb" y ejecute las celdas según sea necesario. Asegúrate de estar utilizando un entorno con GPU en Google Colab durante la ejecución de tu código.

7. Dentro de la carpeta, deberá modificar manualmente los parámetros descritos en el proyecto de investigación. Esto implica cambiar los parámetros (FeatSelect y FS) para el tipo de selección de características utilizado. Estos parámetros se modificarán en el archivo "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/utils/params.py".

8. Además, en el archivo "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/solver.py", debe ajustar los parámetros (N, thresh, a, N2, thresh2, a2) según lo indicado en las tablas de resultados, dependiendo del aerogenerador y el tipo de selección de características.

9. En el archivo "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/scripts/WT.sh", se debe cambiar el umbral (r) de anomalías según se encuentre especificado en las tablas de resultados.

10. Luego, carga los archivos en formato numpy, los cuales resultan de la división de datos, a la carpeta "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/dataset/WT". Estos datos serán fundamentales para el entrenamiento y evaluación del modelo.

10. Finalmente, los resultados del sistema de predicción de fallas se cargarán en la carpeta "/Python Torres Cabrera Joel Stiven Torres/Anomaly-Transformer-modified_Torres-20240418T150717Z-001/Anomaly-Transformer-modified_Torres/Anomaly-Transformer-main/plots/WT".

- Entrenamiento y prueba del modelo TranAD

11. Para entrenar el Modelo TranAD modificado en este estudio, dirígete a la carpeta del modelo llamada "Anomaly-Transformer-modified_Torres-20240418T150717Z-001" y abre el archivo "TranAD_modified_Torres_COLAB 25GB RAM.ipynb". Asegúrate de tener al menos 30 GB de RAM disponible para el entrenamiento; puedes utilizar Google Colab Pro para esto.

12. Además, necesitarás ajustar los parámetros (r , N , thresh , a , N_2 , N_3 , threshEt2 y a_2) según las especificaciones del estudio. Estos parámetros se deben cambiar en el archivo ipynb mencionado.

13. Luego, carga los archivos en formato numpy, los cuales resultan de la división de datos, a la carpeta "/Python Torres Cabrera Joel Stiven Torres/TranAD_Modified_Torres/TranAD/processed/WT_DATA". Estos datos serán fundamentales para el entrenamiento del modelo.

14. Finalmente, los resultados del sistema de predicción de fallas se cargarán en la carpeta "/Python Torres Cabrera Joel Stiven Torres/TranAD_Modified_Torres/TranAD/plots/TranAD_WT_DATA".

Errores que podrías tener al ejecutar el código.

15. En caso de que surja algún error en la ejecución por desactualización, es fundamental asegurarse de actualizar las modificaciones que puedan haber ocurrido en las librerías utilizadas o en la versión de Python en Google Colab. Estos cambios pueden provocar incompatibilidades con algunas librerías. Mantener todo actualizado garantiza un funcionamiento óptimo del código y previene posibles errores relacionados con versiones desactualizadas. También puede contactarse con el equipo de responsables de este proyecto.

Anexo 6. Curva de potencia del aerogenerador GOLDWIND GW 70/1500

Curva de potencia

GW 70/1500

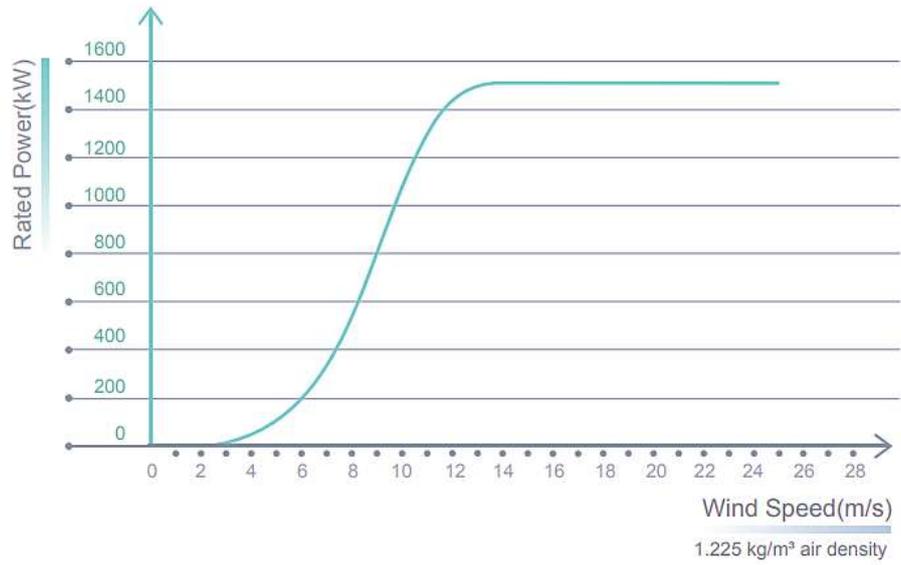


Figura 211. Curva de potencia del aerogenerador GOLDWIND GW/1500

Anexo 7. Acuerdo de confidencialidad

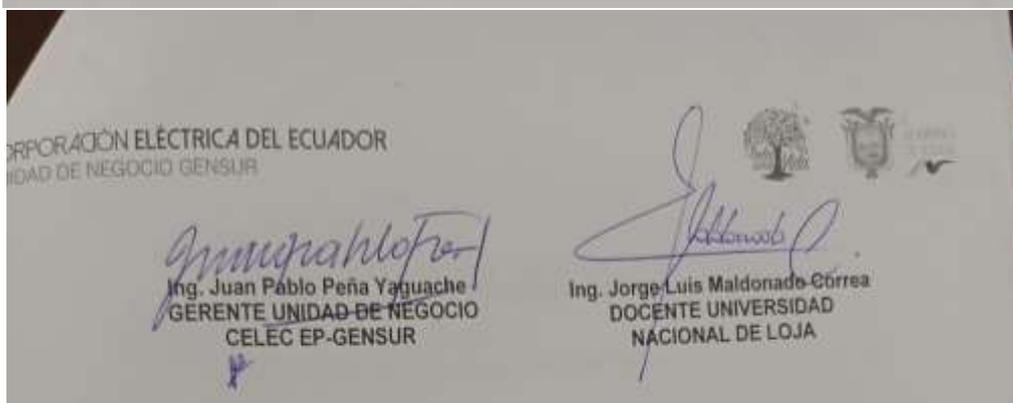
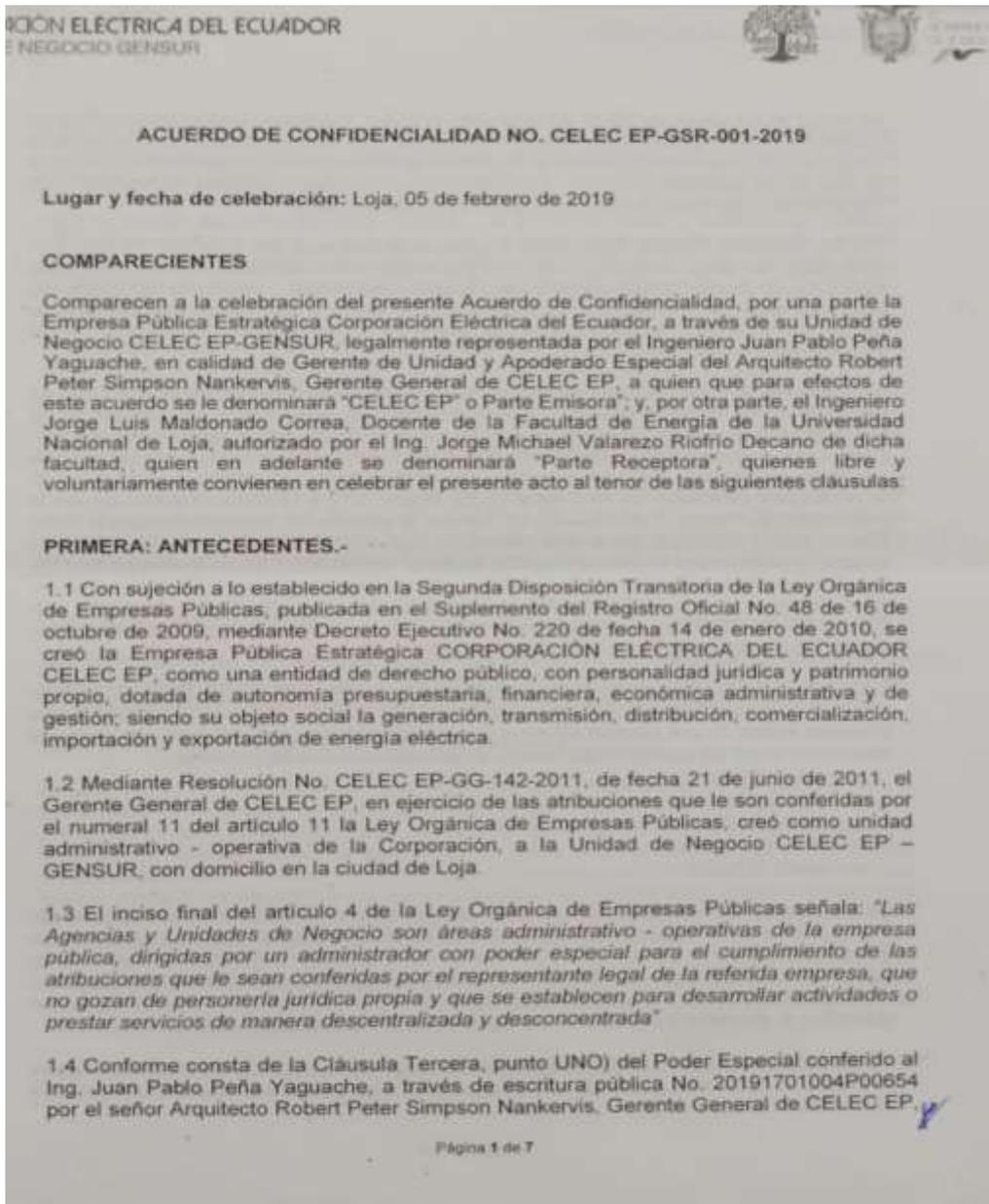


Figura 212. Acuerdo de confidencialidad

Anexo 8. Potencia vs Temperatura IGBT

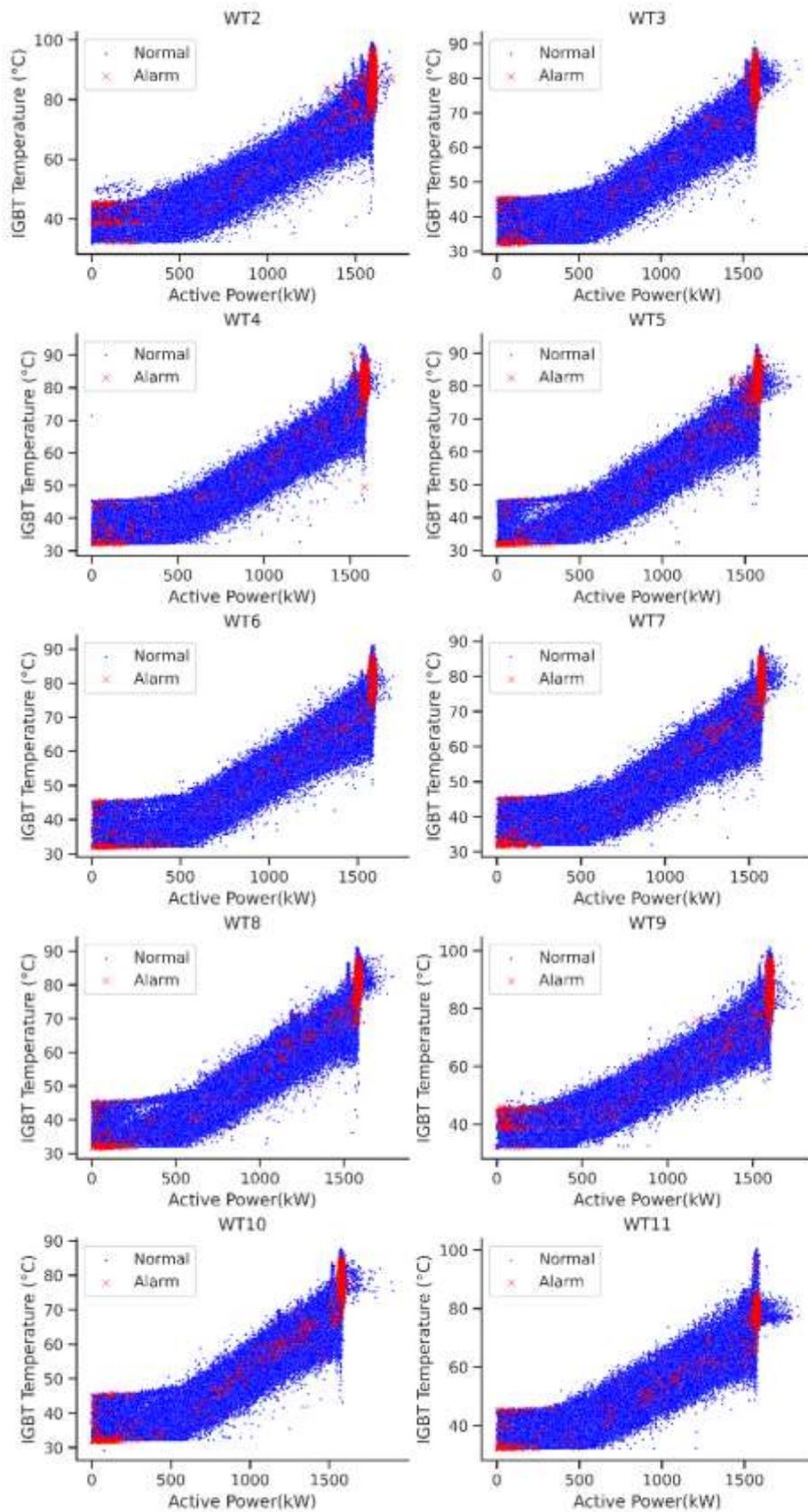


Figura 213. Potencia vs Temperatura IGBT

Anexo 9. Certificado de traducción del Resumen



Mg. Yanina Quizhpe Espinoza
Licenciada en Ciencias de Educación mención Inglés
Magister en Traducción y mediación cultural

Celular: 0989805087
Email: yaniges@icloud.com
Loja, Ecuador 110104

Loja, 07 de marzo 2023

Yo, Lic. Yanina Quizhpe Espinoza, con cédula de identidad 1104337553, docente del Instituto de Idiomas de la Universidad Nacional de Loja, y certificada como traductora e interprete en la Senescyt y en el Ministerio de trabajo del Ecuador con registro MDT-3104-CCL-252640, certifico:

Que tengo el conocimiento y dominio de los idiomas español e inglés y que la traducción del resumen de trabajo de integración curricular **Predicción de fallas en los aerogeneradores de la Central Eólica Villonaco, aplicando la técnica de aprendizaje profundo basado en transformadores**, cuya autoría del estudiante Joel Stiven Torres Cabrera, con cédula 1150343463, es verdadero y correcto a mi mejor saber y entender.

Atentamente

YANINA
BELEN
QUIZHPE
ESPINOZA

Firmado digitalmente por
YANINA BELEN
QUIZHPE
ESPINOZA
Fecha: 2023.03.07
0944:01:05:00

Yanina Quizhpe Espinoza.

Traductora

Full text translator: servicios de traducción

Figura 214. Certificado de traducción del Resumen