

## Modelo de Diagnóstico Predictivo con Inteligencia Artificial para Fallas en Interruptores de Alta Tensión

Adriana Carolina Galindo Vargas, Gonzalo Fandiño Olaya  
E.mail: agalindo@intercolombia.com – gfandino@gmail.com  
Medellín – Colombia

### Resumen

En este estudio, se desarrolló un modelo de inteligencia artificial (IA) para el diagnóstico de fallas en interruptores de alta tensión en sistemas eléctricos de transmisión de 230 kV, con el objetivo de optimizar la identificación de patrones y tendencias de falla en componentes eléctricos asociados a interruptores de alta tensión para facilitar el registro de eventos operativos.

Utilizando técnicas de aprendizaje automático, como clasificador de árboles extra, árboles de decisión y Random Forest, se analizaron datos históricos de fallas en mantenimiento. Para abordar el problema de desbalance de datos en fallas, se aplicó la técnica de sobremuestreo SMOTE. El modelo Random Forest logró una precisión del 99.7% en el diagnóstico de fallas en interruptores, optimizando los tiempos de respuesta y reduciendo falsos positivos en un 15%. Este sistema puede integrarse en estrategias de mantenimiento predictivo en el sector eléctrico, permitiendo detectar fallas con antelación y mejorar la seguridad operativa.

**Palabras clave:** *Inteligencia Artificial, Diagnóstico de Fallas, Mantenimiento Predictivo, Random Forest, Alta Tensión, Interruptor, SCADA.*

### Introducción

La inteligencia artificial (IA) está revolucionando la gestión del mantenimiento de activos, mediante la implementación de metodologías de analítica avanzada como algoritmos de aprendizaje automático para el análisis de datos históricos y en

tiempo real de los equipos y activos operativos del sistema eléctrico, permitiendo prever fallas potenciales antes de que ocurran y definir acciones de mantenimiento optimizadas.

Los modelos de mantenimiento tradicionales con un enfoque preventivo o correctivo ya no son suficientes para la planeación de los mantenimientos, por lo cual la implementación de mantenimientos predictivos, de diagnóstico y prescriptivos son claves para alcanzar los objetivos de disponibilidad, confiabilidad, reducción de costos y la calidad del servicio (IDB, 2023). Es por esto que la inteligencia artificial y los modelos de analítica avanzada en el mantenimiento permiten la toma de decisiones inteligentes y son los pilares de la transformación digital en las organizaciones del sector eléctrico.

En Sistemas de Transmisión Eléctrica de alta tensión, los interruptores desempeñan un papel clave en la protección de infraestructuras, la seguridad de las personas y la disponibilidad del servicio eléctrico. Las fallas en interruptores de alta tensión generan interrupciones en el servicio eléctrico, riesgos para la seguridad y costos elevados, lo que resalta la necesidad de estrategias predictivas efectivas. La identificación de fallas en interruptores es un desafío debido a las diferentes variables que están asociadas a estos equipos como el tipo, los modos de falla, los fabricantes, las características técnicas por familia de equipos, la ubicación geográfica, las condiciones ambientales y operativas que inciden en los equipos.

Los estudios de confiabilidad y análisis de fallas en interruptores generalmente utilizan análisis estadísticos con el fin de tomar mejores decisiones

para optimizar la gestión del mantenimiento. (Gondres Torné, 2018)

El presente estudio se enfoca en diseñar un modelo de diagnóstico basado en IA para la identificación de fallas en interruptores de alta tensión, reduciendo el tiempo de diagnóstico y permitiendo una gestión más eficiente del mantenimiento. Nuestro estudio aplica un enfoque de aprendizaje automático con modelos avanzados como Random Forest para lograr un diagnóstico proactivo. El objetivo general del trabajo es desarrollar un modelo de IA para diagnóstico de fallas en interruptores de alta tensión mediante el análisis de las variables eléctricas del sistema durante las fallas, que permita reducir tiempos de respuesta y optimizar la identificación del componente fallado. Los objetivos específicos incluyen la exploración y análisis de datos históricos de eventos operativos para identificar patrones y tendencias de fallas en los interruptores, la implementación y selección de algoritmos para el desarrollo y evaluación del modelo.

### Metodología

Para el desarrollo de este trabajo se siguieron los pasos definidos en el modelo de Cross Industry Standard Process for Data Mining (CRISP-DM), utilizado para orientar proyectos de minería de datos (IBM, 2021).

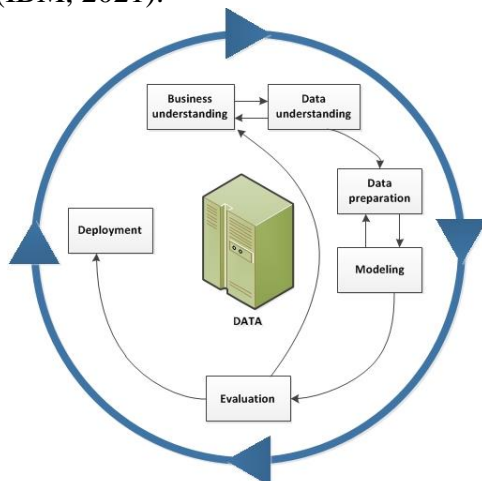


Fig 1. Ciclo de vida de minería de datos.

En primer lugar, para la comprensión del negocio, se definieron los objetivos del estudio y se identificó la necesidad en el proceso de la organización para la identificación de fallas en interruptores de alta tensión.

En segundo lugar, para la comprensión de los datos, fue necesario recopilar datos iniciales, se realizó la obtención, limpieza y preprocesamiento de los datos. Se consultaron bases de datos almacenadas en sistemas como SAP y plataformas propias de la empresa, como el Sistema de información de Gestión Operativa, las cuales almacenan información relevante acerca de las condiciones operativas de los equipos, información histórica de salidas de activos y eventos de fallas durante los últimos 10 años, desde enero del 2013 hasta agosto del 2024. Las bases de datos utilizadas son:

- Reporte de condiciones operativas de eventos con datos de variables eléctricas como tensiones [kV], corrientes [kA] y potencia [MW] de apertura de interruptores en el momento de la falla.

- Reporte de salida de activos con históricos de eventos con información de activos operativos fallados, fecha, hora, tipo, causa y origen del evento.

Estos reportes se descargaron en hojas de cálculo y para la limpieza y preprocesamiento de los datos se trabajaron como archivos tipo CSV.

En tercer lugar, para la preparación de los datos, se seleccionaron, limpiaron y transformaron los datos, identificando las causas más comunes de componentes de fallas en el sistema y se identificaron las correlaciones entre variables.

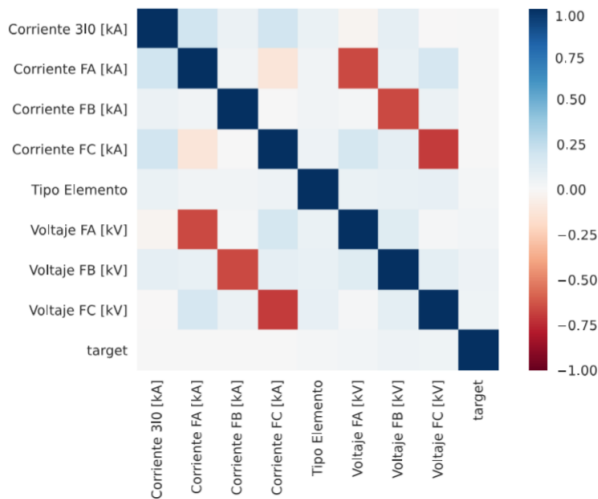


Fig 2. Matriz de correlación.

Como se observa en la Figura 2, se encontraron altas correlaciones entre los voltajes al momento de las fallas, pero se decidió conservar los datos dado que estos valores son útiles para la identificación de las fallas por fase.

Después de una limpieza de datos y consolidación, se generó un conjunto de datos balanceado utilizando la técnica de SMOTE (Synthetic Minority Over-sampling Technique) para mejorar la capacidad predictiva de los modelos en condiciones de datos desbalanceados.

Posteriormente, se seleccionaron y aplicaron técnicas de modelado, ajustando los parámetros para optimizar los resultados del modelo. El análisis de datos y el entrenamiento de modelos se realizaron en Python, utilizando bibliotecas como PyCaret y SKlearn para el desarrollo de algoritmos de clasificación, Pandas y NumPy para la manipulación y procesamiento de datos.

Para el desarrollo del modelo, mediante PyCaret se evaluaron y compararon diferentes modelos de los cuales seleccionamos los siguientes tres modelos con mejor desempeño:

**Random Forest:** Este es un modelo avanzado que combina múltiples árboles de decisión para obtener una clasificación más precisa y robusta. Al utilizar un conjunto de árboles, es capaz de reducir el riesgo de sobreajuste y mejorar la precisión en la identificación de fallas, especialmente en condiciones de datos variados.

**Clasificador de árboles extra:** Al igual que el algoritmo de bosques aleatorios, este algoritmo crea numerosos árboles de decisión, pero el muestreo de cada árbol es aleatorio, sin sustitución. Con ello se crea un dataset para cada árbol con muestras únicas

**Árboles de decisión:** Este modelo clasifica los datos dividiendo el conjunto en múltiples ramas con base en características específicas, lo cual permite identificar patrones de fallas en función de variables concretas. Es intuitivo y fácil de interpretar, y resulta eficaz para detectar relaciones importantes en los datos.

A continuación, se evaluaron los modelos empleando métricas como precisión, recall, f1-score y área bajo la curva ROC (AUC) para evaluar la efectividad de cada modelo y seleccionar el mejor.

Finalmente, se guardaron los modelos y utilizando las herramientas MLflow y Streamlit se realizó un despliegue de prueba del modelo seleccionado.

## Resultados

Para el desarrollo del modelo, mediante PyCaret se evaluaron y compararon diferentes modelos de los cuales seleccionamos los tres modelos con mejor desempeño:

Model		Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
rf	Random Forest Classifier	0.9993	0.9875	0.9233	0.9185	0.9206	0.9202	0.9204	3.3890
et	Extra Trees Classifier	0.9992	0.9993	0.9305	0.8821	0.9034	0.9029	0.9044	2.6260
dt	Decision Tree Classifier	0.9988	0.9648	0.9305	0.8116	0.8647	0.8641	0.8673	2.1640
lightgbm	Light Gradient Boosting Machine	0.9966	0.9880	0.9095	0.5602	0.6912	0.6896	0.7112	2.3680
dummy	Dummy Classifier	0.9959	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	2.0990
knn	K Neighbors Classifier	0.9883	0.9704	0.9438	0.2613	0.4068	0.4030	0.4912	2.5450
gbc	Gradient Boosting Classifier	0.9328	0.9724	0.9038	0.0533	0.1006	0.0936	0.2102	3.8650
ada	Ada Boost Classifier	0.8437	0.9178	0.8543	0.0222	0.0433	0.0355	0.1224	2.4810
ridge	Ridge Classifier	0.7317	0.6223	0.3786	0.0058	0.0114	0.0033	0.0161	2.2790
lda	Linear Discriminant Analysis	0.7317	0.6223	0.3786	0.0058	0.0114	0.0033	0.0161	2.0240
svm	SVM - Linear Kernel	0.6671	0.5814	0.4205	0.0036	0.0071	0.0004	0.0128	2.6070
lr	Logistic Regression	0.5432	0.6066	0.6252	0.0056	0.0112	0.0030	0.0218	3.3200
qda	Quadratic Discriminant Analysis	0.2228	0.6046	0.9171	0.0048	0.0096	0.0014	0.0213	2.0720
nb	Naive Bayes	0.1259	0.5253	0.9248	0.0043	0.0086	0.0004	0.0093	2.2020

Tabla I. Comparación de modelos

El modelo Random Forest: obtiene el mejor desempeño en las principales métricas, exceptuando AUC y Recall. Esto indica una alta precisión general, un buen equilibrio identificando las instancias positivas y minimizando los falsos positivos. Los tiempos de entrenamiento fueron los más altos.

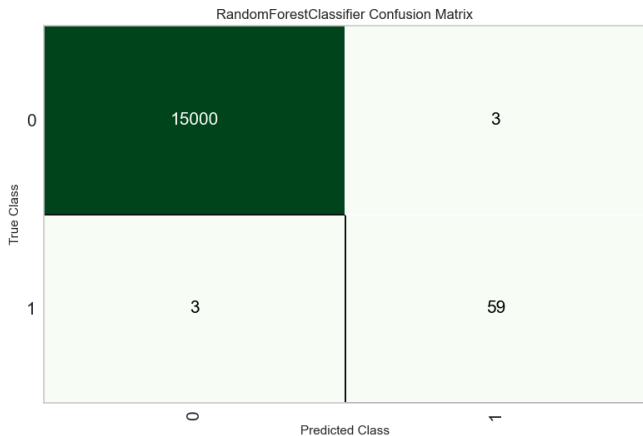


Fig 3. Matriz de confusión para modelo de Random Forest

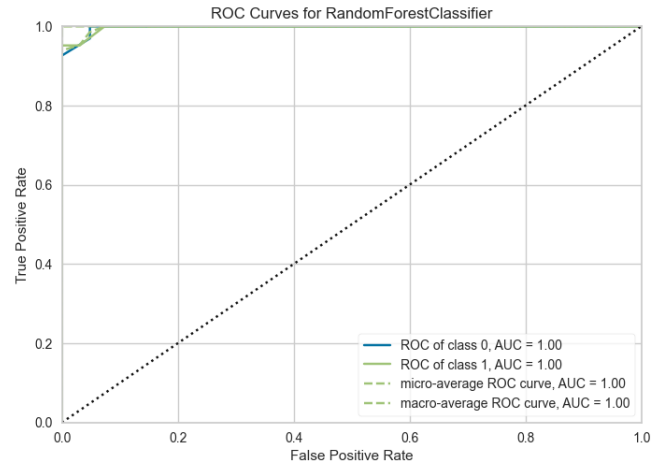


Fig 4 . Curva ROC para el modelo de Random Forest.

El modelo Clasificador de árboles extra: es una alternativa eficiente en comparación con el modelo de Random Forest, con un AUC ligeramente más alto, pero con tiempos de entrenamiento menores. Esto lo convierte en una opción atractiva para proyectos con grandes conjuntos de datos o restricciones de tiempo.

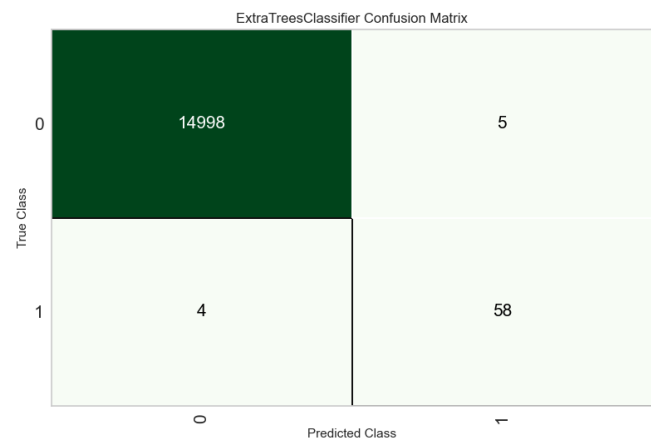


Fig 5. Matriz de confusión para el modelo Clasificador de árboles extra.

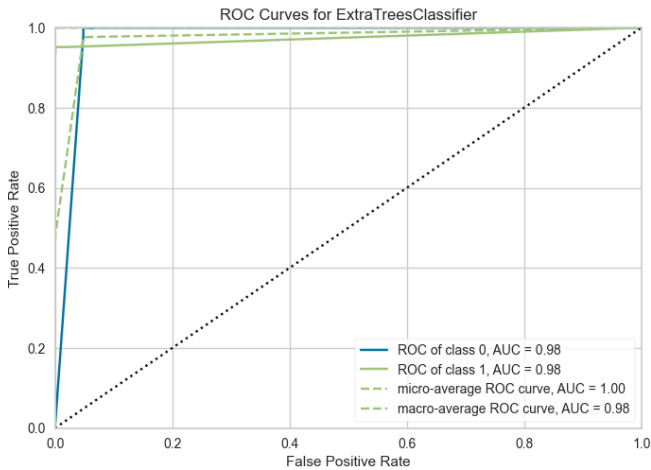


Fig 6. Curva ROC para el modelo Clasificador de árboles extra.

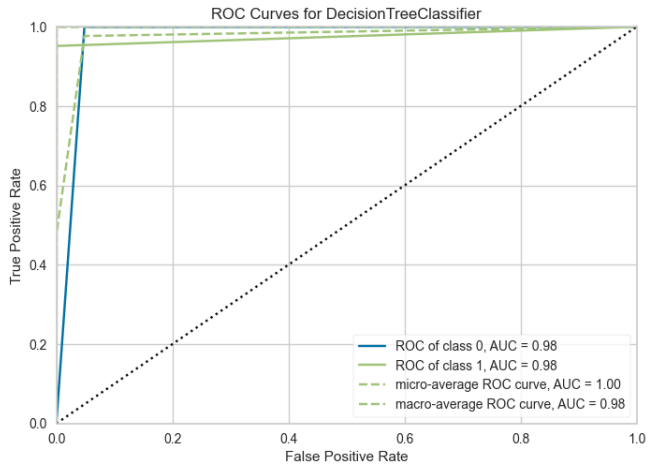


Fig 8. Curva ROC para modelo de árboles de decisión.

El modelo árboles de decisión: obtuvo un buen desempeño en general, pero es inferior en comparación a los modelos Random Forest y Clasificador de árboles extra. El AUC es notablemente más bajo, lo que sugiere que puede tener dificultades para distinguir entre las clases en algunos casos. El tiempo de entrenamiento es el más bajo de los tres modelos.

Modelo y despliegue: De acuerdo con los resultados y utilizando las herramientas MLflow y Streamlit se realizó el despliegue de prueba con la implementación del modelo Random Forest, con lo cual se busca una mayor interacción con el usuario y que este algoritmo sea incorporado al proceso de identificación de fallas de forma automática.

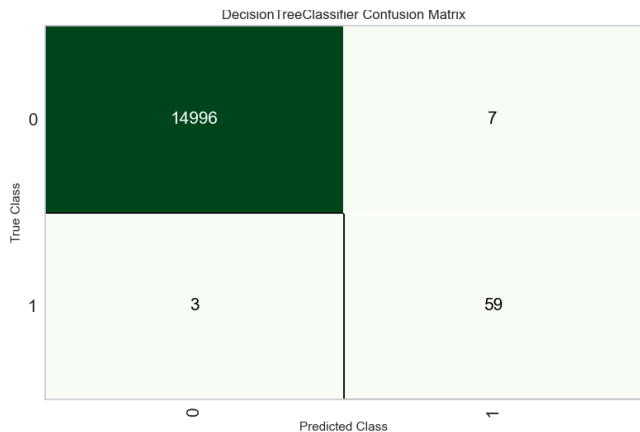


Fig 7. Matriz de confusión para modelo de árboles de decisión.

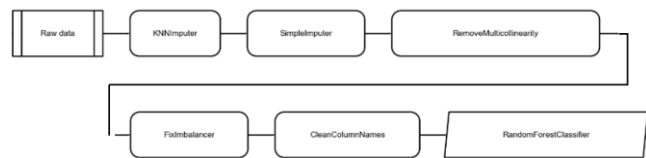


Fig 9. Pipeline del modelo de diagnóstico



### Modelo Diagnóstico Falla Interruptor



Fig 10. Despliegue del modelo de diagnóstico

### Discusión

Los resultados obtenidos indican que los modelos de IA, particularmente Random Forest, son herramientas efectivas para el diagnóstico e identificación de fallas en interruptores de alta tensión. El excelente rendimiento del modelo Random Forest en la clasificación precisa de eventos de falla se debe a su capacidad para manejar datos desbalanceados y aprovechar múltiples árboles de decisión, reduciendo la posibilidad de errores de clasificación.

Este estudio contribuye a la gestión del mantenimiento en sistemas eléctricos, alineándose con los objetivos organizacionales de la compañía hacia la transformación digital y la implementación de modelos de analítica en el mantenimiento que permitan la toma de decisiones inteligentes

En el corto plazo se plantea desarrollar modelos analíticos que permitan identificar la salud de los interruptores de acuerdo con su número de operaciones y magnitudes de corrientes despejadas.

En el mediano plazo, se plantea integrar este modelo a los sistemas tipo SCADA, de modo que se pueda realizar un monitoreo continuo de los

equipos y la predicción de fallas antes de que ocurran, optimizando así los índices de salud, vida útil de los equipos, reduciendo costos y mejorando la disponibilidad de los equipos.

Estos desarrollos enfrentan ciertas limitaciones, al tener una alta dependencia a la trazabilidad y calidad de los datos.

### Conclusiones

En conclusión, la aplicación de IA para la detección de fallas en interruptores de alta tensión no sólo mejora la eficiencia operativa, sino que también contribuye a la reducción de tiempos de diagnóstico.

El modelo Random Forest es una herramienta robusta que puede integrarse a sistemas SCADA para monitoreo continuo, mejorando la confiabilidad de los equipos y optimizando los costos operativos."

La robustez y capacidad de generalización del modelo Random Forest lo convierte en el modelo más preciso para esta aplicación. Comparado con los otros modelos, Random Forest evidenció mayor robustez, lo cual es fundamental en aplicaciones de diagnóstico donde la precisión y confiabilidad son esenciales para evitar interrupciones del servicio.

El modelo Clasificador de árboles extra es una alternativa eficiente en comparación con el modelo de Random Forest, pero con tiempos de entrenamiento menores. Esto lo convierte en una opción atractiva para proyectos con grandes conjuntos de datos o restricciones de tiempo.

El modelo de árbol de decisión es adecuado para la identificación de patrones claros en los datos, aunque podría ser susceptible a problemas de sobreajuste.

## Referencias

[ 1 ] Gondres Torné, I. L. (2018). Gestión del mantenimiento a interruptores de potencia. Estado del arte. Ingeniare. Revista chilena de ingeniería, 26(2), 192-202. <https://dx.doi.org/10.4067/S0718-33052018>. [En línea].

[ 2 ] IBM. (2021). Obtenido de Conceptos básicos de ayuda de CRISP-DM: <https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview>[En línea].

[ 3 ] IDB. (Mayo de 2023). Tecnologías de Inteligencia Artificial (AI) en el Mantenimiento de Activos del Sector. Obtenido de <https://colombiainteligente.org/tendencias/tecnologias-de-inteligencia-artificial-ai-en-el-mantenimiento-de-activos-del-sector-electrico/>[En línea].

Carolina Galindo es analista de gestión de mantenimiento en ISA INTERCOLOMBIA, ingeniera electricista, especialista en Big Data e Inteligencia de Negocios de la Universidad EIA y Magister en Gerencia de Proyectos de la Universidad EAFIT, con 10 años de experiencia en el sector de transmisión de energía. Póngase en contacto con ella en [agalindo@intercolombia.com](mailto:agalindo@intercolombia.com)

Gonzalo Fandiño es ingeniero Especialista en HMV Ingenieros, ingeniero electricista de la Universidad Nacional de Colombia, especialista en Big Data e Inteligencia de Negocios de la Universidad EIA y en Gestión Energética y Ambiental de la Universidad de la Salle. Actualmente interesado en desarrollo y aplicación de modelos de Machine Learning en la industria Eléctrica. Póngase en contacto con él en [gfandino@gmail.com](mailto:gfandino@gmail.com).

## NOTA

Adriana Carolina Galindo Vargas  
Celular 3188224364  
Calle 12 Sur # 18 – 168, Medellín – Colombia

Gonzalo Fandiño Olaya  
Celular 3166269705  
Calle 70 #7-30, Bogotá – Colombia