



UNIVERSIDAD POLITÉCNICA SALESIANA

SEDE QUITO

CARRERA DE ELECTRICIDAD

**DESARROLLO DE UN MODELO ESTADÍSTICO
BASADO EN LA CORRELACIÓN DE PEARSON
PARA ANTICIPAR CONDICIONES
PRECURSORAS DE FALLOS ELÉCTRICOS EN
UN SISTEMA DE DISTRIBUCIÓN DE PRUEBA
REPRESENTATIVO**

Trabajo de titulación previo a la obtención del
Título de Ingeniero Eléctrico

AUTOR: JEOSABETH LIZBETH MOROMENACHO

COLUMBA

TUTOR: CARLOS ANDRÉS BARRERA SINGAÑA

Quito -Ecuador

2026

**CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE
TITULACIÓN**

Yo, Jeosabeth Lizbeth Moromenacho Columba con documento de identificación N°1751143031 manifiesto que:

Soy el autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Quito, 14 de enero del año 2026

Atentamente,

A handwritten signature in blue ink, appearing to read 'Liz M', is written above a horizontal dashed line.

Jeosabeth Lizbeth Moromenacho Columba
1751143031


**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE
TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Jeosabeth Lizbeth Moromenacho Columba con documento de identificación No. 1751143031, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor del Artículo Académico “Desarrollo de un modelo estadístico basado en la correlación de Pearson para anticipar condiciones precursoras de fallos eléctricos en un sistema de distribución de prueba representativo”, el cual ha sido desarrollado para optar por el título de: Ingeniero Eléctrico, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Quito, 14 de enero del año 2026

Atentamente,



Jeosabeth Lizbeth Moromenacho Columba
1751143031

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Carlos Andrés Barrera Singaña con documento de identificación N° 0503503336, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: DESARROLLO DE UN MODELO ESTADÍSTICO BASADO EN LA CORRELACIÓN DE PEARSON PARA ANTICIPAR CONDICIONES PRECURSORAS DE FALLOS ELÉCTRICOS EN UN SISTEMA DE DISTRIBUCIÓN DE PRUEBA REPRESENTATIVO, realizado por Jeosabeth Lizbeth Moromenacho Columba con documento de identificación N° 1751143031, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Quito, 14 de enero del año 2026

Atentamente,



Ing. Carlos Andrés Barrera Singaña, MSc

0503503336

ÍNDICE GENERAL

1	Introducción	2
2	Marco teórico.....	4
2.1	Sistemas de distribución y fallos eléctricos	4
2.1.1	Fallos eléctricos en sistemas de distribución	5
2.2	Variables eléctricas como indicadores de fallos	5
2.3	Modelos estadísticos aplicados a sistemas eléctricos.....	5
2.4	Correlación de Pearson	6
3	Planteamiento del problema y solución al problema	6
3.1	Modelado del sistema IEEE de 33 barras.....	6
3.2	Definición de escenarios operativos y de fallo	8
3.3	Recolección y procesamiento de datos.....	9
3.4	Análisis estadístico mediante correlación de Pearson	10
3.4.1	Definición de la variable objetivo.....	10
3.4.2	Coeficiente de correlación de Pearson	10
3.4.3	Matriz de correlación.....	10
3.4.4	Selección de variables relevantes	10
3.4.5	Construcción del conjunto de variables del modelo.....	10
3.5	Implementación del Modelo de Predicción de Fallos.....	11
3.5.1	Selección de variables de entrada.....	11
3.5.2	Definición de las variables del modelo.....	11
3.5.3	División del conjunto de datos.....	11
3.5.4	Formulación del modelo de regresión logística	11
3.5.5	Entrenamiento del modelo	12
3.5.6	Evaluación del modelo.....	12
3.6	Validación del Modelo	12
3.6.1	Matriz de confusión	12
3.6.2	Exactitud (Accuracy)	12
3.6.3	Precisión (Precision)	12
3.6.4	Sensibilidad o Tasa de Verdaderos Positivos (Recall).....	12
3.6.5	Medida F1 (F1-score)	12
3.6.6	Curva ROC y Área Bajo la Curva (AUC)	13
4	Análisis de Resultados	13

4.1	Variables Eléctricas del Sistema Base	13
4.2	Variables del Sistema con Fallos Eléctricos	15
4.3	Resultados de Correlación de Pearson.....	18
4.4	Métricas del Modelo de Predicción Planteado	19
5	Discusión de Resultados	21
6	Conclusiones	22
7	Trabajos Futuros.....	23
8	Referencias.....	23
8.1	Matriz de Estado del Arte	27
8.2	Resumen de Indicadores	30

ÍNDICE DE FIGURAS

Figura 1. Sistemas de Distribución Eléctrica [23].....	4
Figura 2. Fallos en Sistemas Eléctricos de Potencia [26].....	5
Figura 3. Esquema del Modelo IEEE 33 Barras [35].....	7
Figura 4. Esquema del Sistema IEEE 33 Barras en Python - OpenDSS.....	7
Figura 5. Ejemplo de Generación de Fallas en el Sistema IEEE 33 Barras	9
Figura 6. Perfil de Voltaje del Sistema IEEE Simulado en Estado Normal	13
Figura 7. Corrientes de Línea del Sistema IEEE Simulado en Estado Normal.....	13
Figura 8. Pérdidas de Potencia del Sistema IEEE Simulado en Estado Normal.....	14
Figura 9. Comparación de Variables Eléctricas con Variación de Cargabilidad del Sistema	14
Figura 10. Esquema de Simulación de Falla en el Nodo 8.....	15
Figura 11. Esquema de Simulación de Falla en el Nodo 18	15
Figura 12. Esquema de Simulación de Falla Monofásica en el Nodo 25.....	16
Figura 13. Generación de Múltiples Fallos en Nodos Aleatorios.....	16
Figura 14. Distribución Estadística de Voltaje Mínimo en los Escenarios Generados	17
Figura 15. Relación corriente–voltaje mínimo.....	18
Figura 16. Relación entre pérdidas técnicas y potencia total.....	18
Figura 17. Diagrama de Calor de la Matriz de Correlación de Pearson.....	19
Figura 18. Curva ROC de Detección de Fallos	20
Figura 19. Curva de Precisión del Modelo Planteado	21
Figura 20. Distribución de Probabilidades del Modelo	21
Figura 21. Resumen e indicador de la temática – Estado del arte.....	30
Figura 22. Indicador de formulación del problema - Estado del arte.	30
Figura 23. Indicadores de Solución Propuesta – Estado del Arte.....	30

ÍNDICE DE TABLAS

Tabla 1. Características del Sistema IEEE 33 Barras	7
Tabla 2. Variación de la Demanda en el Sistema Planteado	8
Tabla 3. Tipos de Fallas Simuladas	8
Tabla 4. Variables eléctricas utilizadas para la construcción de la data	9
Tabla 5. Estadísticas de Voltaje en Condiciones Normales del Sistema IEEE 33 Barras	13
Tabla 6. Estadísticas de Voltaje en Condiciones de Falla en el Nodo 8	15
Tabla 7. Estadísticas de Voltaje en Condiciones de Falla en el Nodo 18	15
Tabla 8. Esquema de Simulación de Falla en el Nodo 25	16
Tabla 9. Estadísticas Eléctricas del Sistema en Múltiples Escenarios de Fallos	17
Tabla 10. Resultados de Correlación de Pearson más Representativos	19
Tabla 11. Resultados de Evaluación del Modelo de Predicción	20
Tabla 12. Matriz de Confusión del Sistema de Clasificación	20
Tabla 13: Matriz de estado del arte.	27

DESARROLLO DE UN MODELO ESTADÍSTICO BASADO EN LA CORRELACIÓN DE PEARSON PARA ANTICIPAR CONDICIONES PRECURSORAS DE FALLOS ELÉCTRICOS EN UN SISTEMA DE DISTRIBUCIÓN DE PRUEBA REPRESENTATIVO

Resumen

Al hablar del crecimiento continuo de la demanda eléctrica y el aumento de la cargabilidad han incrementado la probabilidad de fallos en los sistemas de distribución eléctrica. En el presente estudio se desarrolló un modelo estadístico basado en la correlación de Pearson para anticipar fallos eléctricos en el sistema de prueba IEEE de 33 barras. Se modeló el sistema dentro de OpenDSS y se simularon diversos escenarios operativos y de fallo para generar una base de datos eléctricos. Posteriormente, se analizaron las variables eléctricas registradas para aplicar la correlación de Pearson seleccionando las que mayor relación estadística presentan con la ocurrencia de fallos, en donde la mayor correlación positiva se encontró en la corriente máxima con 0.1646 y las pérdidas activas con 0.1645. Se implementó un modelo de regresión logística para evaluar la capacidad de anticipación de fallos, cuyos resultados presentaron un recall del 100 % para la clase de fallo, una curva ROC con un AUC de 0.40 y superposición de probabilidades de detección lo cual demuestra que el modelo tiene una limitada capacidad de diferenciación entre ciertas condiciones operativas pero demostrando que esta metodología es factible como herramienta complementaria para la planificación del mantenimiento preventivo.

Palabras clave: Probabilidad de fallos, Correlación de Pearson, IEEE de 33 barras, Modelo de regresión logística, Mantenimiento preventivo.

Abstract

The continuous growth in electricity demand and the increase in load capacity have increased the likelihood of failures in electrical distribution systems. In this study, a statistical model based on Pearson's correlation was developed to predict electrical failures in the IEEE 33-bus test system. The system was modeled within OpenDSS, and various operational and failure scenarios were simulated to generate an electrical database. Subsequently, the recorded electrical variables were analyzed to apply Pearson's correlation, selecting those with the strongest statistical relationship to the occurrence of failures. The highest positive correlation was found in the maximum current (0.1646) and active losses (0.1645). A logistic regression model was implemented to evaluate the model's ability to anticipate failures. The results showed a 100% recall for the failure class, an ROC curve with an AUC of 0.40, and overlapping detection probabilities. This demonstrates that the model has a limited ability to differentiate between certain operating conditions but shows that this methodology is feasible as a complementary tool for preventive maintenance planning.

Keywords: Probability of failures, Pearson correlation, IEEE 33-bar, Logistic regression model, Preventive maintenance.

1 Introducción

El crecimiento exponencial de la demanda, junto con el aumento de la digitalización y la integración de fuentes distribuidas, ha aumentado la operatividad de los sistemas eléctricos de distribución (SED) [1] [2]. Los fallos eléctricos en estos SED representan una proporción substancial relacionada a la continuidad de los servicios, la integridad de los elementos y los gastos respecto al funcional de los fieles [3]. Por lo tanto, la identificación de las condiciones que provienen de estas tragedias es esencial para planificar y llevar a cabo medidas activas y preventivas en el mantenimiento [4].

La falla monofásica a tierra se presenta como una de las más frecuentes en los SED. Este error ocurre cuando uno de los conductores de fase se coloca en contacto con tierra o con partes conductoras conectadas a tierra [5]. Esta condición genera una gran caída de voltaje de la fase en cuestión y un considerable aumento de la corriente de secuencia cero. Es esencial conducir un análisis de tales fallas, ya que, en redes de distribución radiales, como las de zonas periféricas, en particular las rurales, generalmente pasan desapercibidas [6].

Aunque menos común que la falla anterior, la falla trifásica es igualmente muy destructiva debido a la alta corriente de cortocircuito que causa. Este fallo implica que las 3 fases del sistema entran en contacto al mismo tiempo y, como resultado, ocurre la caída simultánea de los voltajes de fase y el aumento de la corriente en todos los cables [7] [8].

Por otro lado, el fallo de sobrecarga se produce debido a una demanda excesiva en comparación con la capacidad del sistema. Esto resulta en un aumento de la potencia aparente, pérdidas resistivas y una disminución gradual de los niveles de voltaje. Se indica que a un análisis del sistema IEEE de 33 barras, al aumentar la

carga al 150% del valor base, el voltaje en algunos nodos disminuyó hasta un 25%, mientras que las pérdidas activas y reactivas también aumentaron significativamente. Sin la implementación de mecanismos de compensación, estas condiciones pueden resultar en un colapso de voltaje [9] [10].

En general, las herramientas tradicionales de monitoreo de sistemas son valiosas, tienden a reaccionar frente a un evento anómalo luego de que este haya sucedido. Contrariamente, los enfoques predictivos son capaces de anticipar comportamientos anómalos porque analizan datos operativos históricos y actuales del sistema. Si tales variables, por ejemplo, corriente, voltaje o potencia, se evalúan en condiciones de operación diferentes, pueden dar lugar a patrones que, funcionando de modo correlativo, pueden anticipar fallas eléctricas [11] [12].

Los métodos aplicados para la localización y predicción de fallos en los SED han demostrado ser altamente limitados en múltiples ocasiones, lo que justifica la necesidad de enfoques alternativos más accesibles y robustos. Por un lado, las técnicas basadas en IA (inteligencia artificial) resultan ser prometedoras, pero dependen de grandes volúmenes de datos que deben estar previamente etiquetados, lo que hace que las mismas sean poco aplicables para redes pobremente instrumentadas o cuando los datos históricos son incompletos [13] [14].

Además de la IA, otros métodos tradicionales, como el análisis de impedancia o la tecnología basada en la reflectometría en el dominio del tiempo (TDR), también presentan varias limitaciones. Por ejemplo, cuando se trata del diagnóstico de cables subterráneos, la interpretación de las señales en el reflectograma resulta ser poco concluyente, comparado ante ámbitos reales. En particular, con respecto a los conductores antiguos, así como a los que tienen varias

capas de aislamiento. En tales casos, la señal puede ser de poca ayuda, ya que puede ser confusa, y los algoritmos de TDR no siempre son universales y requieren ajustes manuales o una interpretación subjetiva por parte de un técnico con experiencia [15] [16].

Ante estos métodos que representan limitantes en su desarrollo en el SED, se aplica una alternativa simple, como lo es la correlación de Pearson. Esta se considera como un método estadístico clásico para estimar posibles relaciones lineales entre dos variables operacionales en los sistemas eléctricos [17] [18].

Existe una gran variedad de temáticas con relación a la predicción de fallas en redes eléctricas, descritas a continuación:

Las condiciones precursoras del artículo propuesto por Bartolomeo et al. [19] consistió en analizar fallos eléctricos, específicamente cortocircuitos monofásicos y trifásicos a tierra dentro de un SED de baja voltaje, utilizando como sistemas estudiados la red de prueba CIGRE 18 buses. El trabajo se centró en cuantificar el impacto de la operación previa y desbalanceada de un cortocircuito, agravada por la interacción de la generación distribuida (GD) monofásica presente en 3 buses del sistema. Para ello, aplicaron los métodos de Monte Carlo y posición de falla. Al generar múltiples escenarios, se determinaron valores de voltaje durante la falla entre 0,43 a 1,0513 p.u., así como aumentos temporales de voltaje con valores cercanos a 1,166 p.u. cuando las condiciones cumplen los umbrales de secuencia positiva son menores a 0,85 y de secuencia cero son mayores a 0,40 para clasificar *swells* (hinchamientos de voltaje). Esta información identifica la mayor probabilidad de voltaje asimétrica durante la falla, constituyendo una condición precursora analizada.

Hariharan et al. [13] efectuaron el análisis de precursores de falla a los valores RMS (*Root Mean Square*) de voltaje y corrientes trifásicas anómalos inducidos por fallas y aplicados en una IEEE 9-bus integrado, junto con fuentes de origen fotovoltaico, eólico y de carga en vehículos eléctricos. Cada falla se índice en diversas líneas de transmisión en el programa MATLAB/Simulink, lo que resulta en 12.000 conjuntos de datos completamente distribuidos al final. Para detectar tales condiciones, aplicaron un autoencoder entrenado, el cual fue comparado con otros métodos como KNN, SVM, y *Random Forest*. La precisión del modelo presentado es de 98% incluso en fallas no vistas, pero inducidas aleatoriamente; sin embargo, el K-cercano vecino disminuye a 21%, SVM a 10% y *Random Forest* a 27%. Por lo tanto, estos resultados numéricos confirman que el autoencoder avisa cualquier desviación del patrón normal para garantizar la resistencia de las condiciones precursoras.

Kong y Nian [20] establecieron como condiciona precursora a la variación de las corrientes en líneas de una microrred de corriente continua (DC) mallada para un sistema de distribución de 4 nodos. Además, aplicaron una correlación de Pearson mejorada, para comparar la corriente muestreada con la de referencia en una ventana deslizante. Los resultados apuntan a tiempos de detección de 2,22 ms en promedio y un máximo inferior a 4,5 ms, mientras que la localización demuestra errores iniciales de 0,49% en posición y 1,03% en resistencia de falla, para finalmente alcanzar una precisión experimental de 0,50% en la ubicación del punto de falla.

Lestari et al. [21] analizaron las condiciones precursoras asociadas con la descoordinación de protecciones, como las variaciones en la dirección del flujo de potencia y en los niveles de corriente de cortocircuito con mayor impacto en el

rendimiento de relés, especialmente en el sistema primario formado por 13 relés distribuidos en un SED con GD. También pretende encontrar variaciones estadísticamente significativas en la detección y el ajuste óptimo de dichos precursores mediante una simulación basada en la optimización por la metaheurística *Firefly*, comparada con el método de prueba manual convencional para la coordinación de relés de sobrecorriente. Los resultados evidenciaron una mejora en la selectividad y rapidez de respuesta del sistema ante fallas, donde se observa una reducción de 0,559 s (método manual) a 0,364s (*Firefly*), disminuyendo el tiempo de operación del Relé.

Liu et al. [11] presentaron varias condiciones precursoras antes del fallo, como los errores de medición, una baja observabilidad, el ruido, la topología variable y los valores faltantes (removidos o no reportados por los sensores), mediante un énfasis en los SED y 3 sistemas de prueba diferentes, CIGRE (14-bus) y *Oberrhein* (70-bus y 179-bus). Para conseguir esta detección, aplicaron el modelo GATNE (*General Attributed Multiplex Heterogeneous Network Embedding*) junto con una atención *cross-modal* y restricciones físicas basadas en el flujo de potencia, lo que facilita la identificación de las discrepancias entre las mediciones reales y los resultados físicos esperados. Para el sistema de 14 barras, el RMSE (*Root Mean Square Error*) del voltaje disminuye a $4,32 \times 10^{-3}$ p.u., menos de la mitad de lo que se obtiene con WLS, es decir, $9,53 \times 10^{-3}$ p.u. En las demás redes 70-bus y 179-bus, obtuvieron un RMSE $2,91 \times 10^{-3}$ p.u. y $3,28 \times 10^{-3}$ p.u.; respectivamente. Por último, el error de carga en línea desciende a alrededor de 2,22% – 2,97%. Esto muestra un enfoque más robusto para abordar el ruido, valores faltantes y los errores de medición.

En base a los antecedentes mencionados, se observa la necesidad de comprobar mediante un modelo de Pearson la influencia de las variables eléctricas del sistema IEEE de distribución con 33 barras logrando identificar cuáles son las que mayor correlación poseen y de esta forma conseguir un modelo que permita predecir fallos de manera adecuada.

2 Marco teórico

2.1 Sistemas de distribución y fallos eléctricos

Los sistemas de distribución establecen como el punto final del sistema eléctrico de potencia y cuyo objetivo principal es transportar la energía eléctrica desde las subestaciones de transmisión hasta los últimos usuarios como se visualiza en la **Figura 1**, manteniendo niveles adecuados de voltaje, continuidad y calidad del servicio. Estos sistemas operan de manera general en niveles de baja y medio voltaje cuya topología suele ser radial por su simplicidad operativa y menores costos de implementación y protección [22].



Figura 1. Sistemas de Distribución Eléctrica [23]

Es un sistema de distribución típico conformado por alimentadores, líneas de distribución, transformadores de distribución y cargas, los cuales mantienen una interacción dinámica bajo diversas condiciones operativas. Debido al incremento en la demanda a nivel mundial, la inclusión de cargas no lineales y la variabilidad en el consumo durante las diferentes épocas del año, generan que los sistemas de distribución estén sujetos a fluctuaciones en sus variables eléctricas

tales como voltaje, corriente y potencia, las cuales afectan de forma directa al desempeño y confiabilidad de todo el sistema [24].

2.1.1 Fallos eléctricos en sistemas de distribución

Los fallos eléctricos en los sistemas de distribución son conocidos como eventos anómalos que alteran el funcionamiento nominal de la red, lo cual produce interrupciones del servicio, deterioro de equipos o riesgos para la seguridad en general. Estos fallos se clasifican de acuerdo con su naturaleza, en fallos monofásicos a tierra, bifásicos y trifásicos como se observa en la **Figura 2**, siendo los fallos monofásicos los que mayor recurrencia se presentan en redes de distribución radial [25].

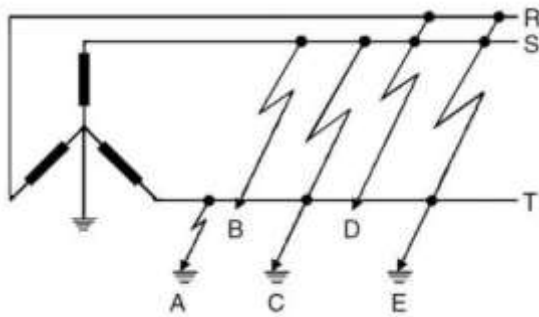


Figura 2. Fallos en Sistemas Eléctricos de Potencia [26]

Las principales causas que producen fallos eléctricos son las sobrecargas prolongadas, envejecimiento del aislamiento, condiciones ambientales como lluvia, humedad, descargas atmosféricas y fallas mecánicas en conductores y equipos. Antes de que se presente un fallo permanente, el sistema eléctrico presenta variaciones progresivas en ciertos parámetros como caídas de voltaje, incrementos de corriente y aumento de pérdidas técnicas, las cuales se consideran como condiciones precursoras del fallo [27].

2.2 Variables eléctricas como indicadores de fallos

Las variables eléctricas de un sistema de distribución son una fuente de información que permite evaluar el estado operativo. Dentro de las variables más relevantes se puede considerar al voltaje en los nodos, la corriente en los alimentadores, la potencia reactiva, activa, pérdidas técnicas y el factor de potencia. Cuando se presentan variaciones anormales en estas variables se pueden mostrar como condiciones precursoras de fallos eléctricos [28].

Un ejemplo son las caídas de voltaje sostenidas las cuales se asocian principalmente a sobrecargas o deterioro en las líneas, y por otro lado, cuando se presentan incrementos significativos en la corriente, esto puede indicar condiciones de estrés térmico en conductores y transformadores. De igual manera, el incremento en las pérdidas técnicas puede evidenciar ineficiencias operativas o problemas iniciales dentro en la red. Cuando se consigue monitorear de manera continua a estas variables se pueden identificar patrones de comportamiento que indiquen cuando se presentan un fallo en el sistema [29].

2.3 Modelos estadísticos aplicados a sistemas eléctricos

El análisis estadístico para sistemas que funcionan de manera eléctrica con potencia es considerado actualmente como una herramienta fundamental para el estudio del comportamiento operativo de las redes, principalmente para los sistemas de distribución, los cuales poseen una alta variabilidad en las condiciones de carga y operación. Se pueden utilizar técnicas estadísticas para procesar grandes volúmenes de datos provenientes de mediciones reales o de simulaciones, con el objetivo de reconocer patrones, tendencias

e identificar relaciones entre variables eléctricas [30].

Los modelos estadísticos que se aplican dentro de la identificación de fallos eléctricos pretenden establecer relaciones entre variables operativas del sistema y la ocurrencia de eventos anómalos. Estos modelos pueden estar basados en técnicas descriptivas, inferenciales o de correlación, y su principal ventaja está en la simplicidad de implementación y la capacidad de interpretación de los resultados [31].

Para un sistema de distribución, la detección temprana de fallos se relaciona con el análisis de múltiples variables eléctricas, tales como niveles de voltaje, corrientes de línea y pérdidas técnicas. Utilizando modelos estadísticos se puede evaluar las variables que tienen mayor influencia o relación con la ocurrencia de fallos [32].

2.4 Correlación de Pearson

Es una medida estadística la cual se encarga de cuantificar el nivel de relación lineal existente entre dos variables cuantitativas. Este coeficiente, comúnmente descrito como r , puede adquirir valores en el intervalo $[-1,1]$, de los cuales los valores que suelen acercarse a 1 corresponden a una correlación efectiva fuerte, mientras que los valores que tienen cercanía a -1 se menciona que existe correlación desfavorable fuerte y sus valores suelen ser próximos a 0 son un indicio de ausencia en la relación entre las variables analizadas [33].

Cuando se trata de análisis de sistemas eléctricos de distribución, la correlación de Pearson puede evaluar cual es la relación existente entre las variables operativas del sistema, tales como voltaje, corriente o pérdidas técnicas, y la ocurrencia de eventos anómalos. Debido a su simplicidad de cálculo y fácil interpretación sirve de

manera efectiva como una herramienta estadística para efectuar estudios exploratorios orientados a la identificación de condiciones que provocan un fallo eléctrico [34].

3 Planteamiento del problema y solución al problema

La metodología propuesta la investigación se basa en un enfoque cuantitativo, exploratorio y aplicado, el cual se basa en la simulación computacional del sistema IEEE de 33 barras. En primera instancia, se realiza el modelado del sistema IEEE dentro del entorno de OpenDSS, tomando en cuenta las características principales y condiciones nominales de este modelo. Posteriormente, se realizan diversos escenarios de operación y de fallos eléctricos para poder generar una base de datos eléctricos que represente el comportamiento del sistema. Por medio de Python se interpretan los datos y se realiza la correlación para verificar las principales relaciones de las variables y la ocurrencia de condiciones de fallos. Finalmente, se generará y evaluará el modelo que permite la predicción de fallos como herramienta de apoyo en la planificación del mantenimiento preventivo en sistemas de distribución eléctrica. A continuación, se presenta a detalle cada uno de los procedimientos realizados para la ejecución del proyecto.

3.1 Modelado del sistema IEEE de 33 barras

El sistema de prueba IEEE de 33 barras es un modelo utilizado para investigaciones relacionadas con sistemas de distribución eléctrica. Este modelo representa una red de distribución radial con un medio voltaje como se observa en la **Figura 3**, el cual está diseñado para evaluar el comportamiento de la red ante diversos escenarios operativos, tales como variaciones de carga, análisis de

pérdidas, reconfiguración de redes y estudios de confiabilidad.

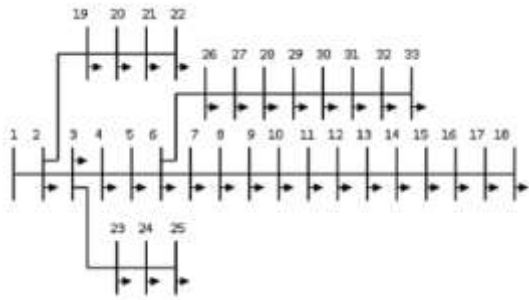


Figura 3. Esquema del Modelo IEEE 33 Barras [35]

Este modelo posee 33 nodos interconectados mediante líneas de distribución, con una única barra de alimentación que representa la subestación principal. Esta red presenta una topología radial, la cual es común en sistemas de distribución reales la cual está configurada con un nivel de voltaje nominal de 12,66 kV. Las cargas están distribuidas a lo largo de los nodos del sistema y se modelan generalmente como cargas constantes de potencia activa y reactiva.

Para el modelado dentro de OpenDSS se utilizan la información que se observa en la **Tabla 1**.

Tabla 1. Características del Sistema IEEE 33 Barras

Característica	Valor
Voltaje Nominal (Fuente)	12,66 kV
Frecuencia de Trabajo (Fuente)	60 Hz
Configuración (Fuente)	Trifásica – Δ (Delta)
Potencia (Fuente)	6 MVA
Cantidad de Barras	33
Cantidad de Cargas	32
Resistencia / Km por Trayecto	[0.01273 0.3864] Ω/km
Inductancia / Km por Trayecto	[0.9337e-3 4.1264e-3] H/km
Capacitancia / Km por Trayecto	[12.74e-9 7.751e-9] F/km

Una vez determinadas las características principales de cada uno de los nodos del sistema, se realiza la implementación dentro de OpenDSS por medio de Python, obteniendo el esquema que se observa en la **Figura 4**.

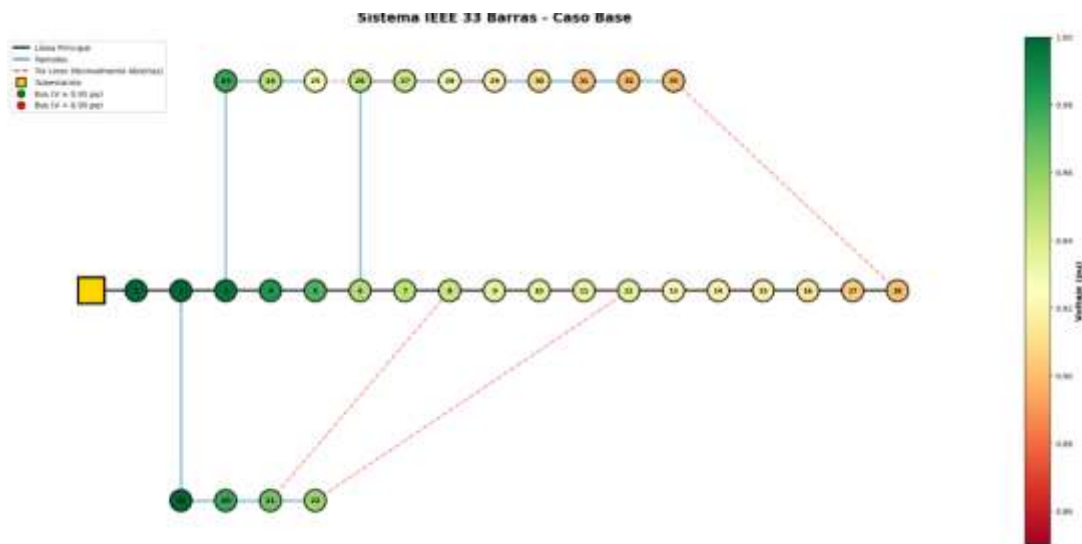


Figura 4. Esquema del Sistema IEEE 33 Barras en Python - OpenDSS

3.2 Definición de escenarios operativos y de fallo

Con el sistema planteado, se procede a determinar la metodología para el planteamiento de fallos. Se establecen los escenarios operativos normales en donde se realiza la variación existente de la demanda eléctrica del sistema, cuyo objetivo es de representar distintos niveles de carga que pueden presentarse en la operación real de una red de distribución. Como se observa en la **Tabla 2**, se consideran incrementos y decrementos porcentuales aplicados a las cargas nominales del sistema IEEE de 33 barras.

Tabla 2. Variación de la Demanda en el Sistema Planteado

Escenario	Nivel de carga	Variación de la demanda	Descripción
E1	Baja	-30 %	Representa periodos de baja demanda, como horas nocturnas o fines de semana
E2	Media	-10 %	Condición de operación cercana al mínimo operativo
E3	Nominal	0 %	Operación normal del sistema
E4	Media-alta	+10 %	Incremento moderado de demanda
E5	Alta	+30 %	Condición de alta carga, representativa de horas pico

Con esto se evalúa el desempeño que tiene el sistema bajo diferentes condiciones de carga realizando la identificación del comportamiento de las variables eléctricas tales como las pérdidas de potencia, el perfil de voltaje y las corrientes en las líneas del sistema sin presentar fallos.

Posteriormente, dentro de la simulación, se ejecutan los tipos de fallos más frecuentes en sistemas de distribución tales como son los fallos monofásicos a tierra y fallos trifásicos como se observa en la **Tabla 3**.

Tabla 3. Tipos de Fallas Simuladas

Tipo de fallo	Descripción	Fases involucradas
Monofásico a tierra	Cortocircuito entre una fase y tierra	Fase A – Tierra
Trifásico	Cortocircuito simultáneo entre las tres fases	A – B – C

Los fallos se simulan en diferentes ubicaciones de las barras del sistema, las cuales se seleccionan de forma aleatoria a lo largo del alimentador principal, con el fin de analizar el impacto de la ubicación del fallo sobre el comportamiento del general del sistema.

Como se visualiza en la **Figura 5**, se genera un fallo en el Nodo 18 del alimentador lo cual permite representar una condición intermedia de falla representando el comportamiento aguas abajo y aguas arriba las variables del sistema ante un fallo eléctrico.

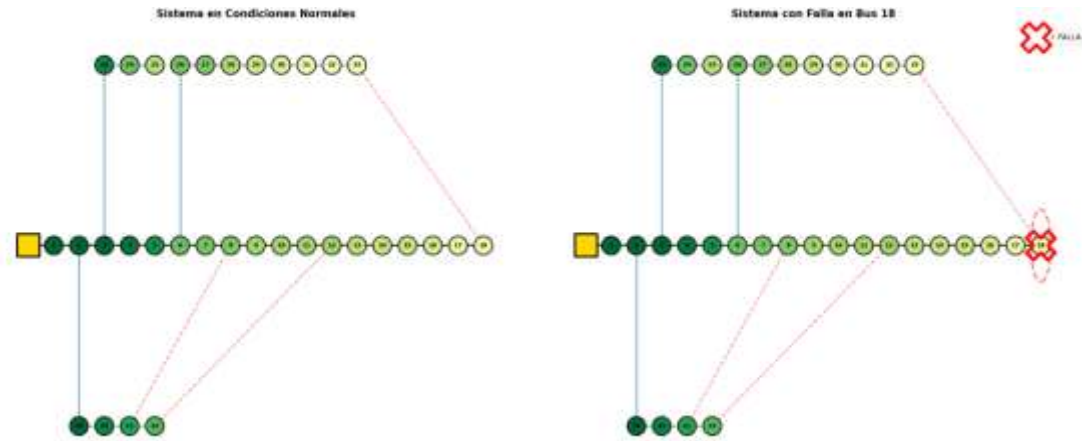


Figura 5. Ejemplo de Generación de Fallas en el Sistema IEEE 33 Barras

3.3 Recolección y procesamiento de datos

En el proceso de recolección de la información, se los obtienen a partir de las simulaciones del sistema IEEE de 33 barras ejecutadas en OpenDSS, considerando las condiciones mencionadas en el punto anterior, las cuales son ejecutadas por medio de instrucciones de Python.

Para garantizar una representatividad estadística de los resultados, se ejecutaron 100 simulaciones en condiciones operativas normales, únicamente presentando variaciones de carga sin considerar la presencia de fallos, mientras que se realizaron 29 simulaciones adicionales bajo condiciones de fallo, ejecutando fallos monofásicos y trifásicos de forma aleatoria en los diferentes nodos del sistema.

Dentro de cada simulación, se registraron en variables eléctricas dentro del entorno de Python las cuales permiten caracterizar el estado actual del sistema en ambos escenarios de prueba. Las variables consideradas para formular la base de datos son voltaje, corriente, potencia y pérdidas técnicas, permitiendo mostrar el comportamiento local en las barras. Como se observa en la **Tabla 4**, se detallan las diferentes variables usadas para que se desarrolle el modelo, en donde se incluyen

las respectivas etiquetas dentro de la programación realizada.

Tabla 4. Variables eléctricas utilizadas para la construcción de la data

Variable	Descripción	Unidad
has_fault	Indicador binario de presencia de fallo (0: operación normal, 1: fallo)	—
max_current_A	Valor máximo de corriente registrado en las líneas del sistema	A
total_current_A	Corriente total agregada del sistema	A
losses_kW	Pérdidas técnicas totales de potencia activa del sistema	kW
total_power_kW	Potencia activa total demandada por el sistema	kW
load_mult	Multiplicador de carga aplicado al escenario simulado	—
avg_voltage_pu	Voltaje promedio del sistema en unidades por unidad	pu
min_voltage_pu	Voltaje mínimo registrado entre todas las barras del sistema	pu
max_voltage_pu	Voltaje máximo registrado entre todas las barras del sistema	p

3.4 Análisis estadístico mediante correlación de Pearson

Por medio del análisis de correlación de Pearson se identifican las conexiones lineales entre variables eléctricas del sistema de distribución y la ocurrencia de fallos eléctricos. Con este enfoque se cuantificará el nivel de asociación entre los parámetros operativos del sistema y los estados de fallos provocados.

3.4.1 Definición de la variable objetivo

En primer lugar, se define una variable binaria denominada *has_fault*, la cual indica si existe o no la presencia de un fallo eléctrico en cada escenario simulado. Esta variable se expresa matemáticamente de la siguiente forma:

$$has_{fault_i} = \begin{cases} 1, & \text{si existe un fallo} \\ & \text{en el escenario } i \\ 0, & \text{si el sistema opera} \\ & \text{en condiciones normales} \end{cases} \quad (1)$$

Esta variable permite evaluar cómo se relacionan las variables obtenidas de la simulación en condiciones normales y cuando se presentan fallos.

3.4.2 Coeficiente de correlación de Pearson

Posteriormente, se calcula el coeficiente de correlación, el cual muestra si existe una relación lineal entre dos variables aleatorias X y Y , definido de la siguiente forma:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (2)$$

donde:

- X_i y Y_i son las observaciones individuales de cada una de las variables eléctricas.
- \bar{X} y \bar{Y} son los valores medios de dichas variables,
- n es el número total de escenarios analizados.

Este coeficiente r se calcula entre cada variable eléctrica del sistema y la variable

has_fault, lo cual cuantificará el grado de asociación entre el comportamiento eléctrico del sistema y la presencia de fallos.

3.4.3 Matriz de correlación

Para desarrollar el análisis de la relación existente entre las variables se construye la matriz de correlación \mathbf{R} , la cual se define como:

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1m} \\ r_{21} & 1 & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & 1 \end{bmatrix} \quad (3)$$

Donde r_{ij} representa el coeficiente de correlación entre las variables i y j , mientras que m es el número total de variables analizadas.

3.4.4 Selección de variables relevantes

Una vez obtenida la matriz de la correlación se especifica el vector de correlación asociado a la variable *has_fault*, expresado como:

$$\mathbf{r}_{fault} = [r_{X_1, has_{fault}}, r_{X_2, has_{fault}}, \dots, r_{X_m, has_{fault}}] \quad (4)$$

Esto se realiza para poder seleccionar para el modelo únicamente las variables más relevantes, para lo cual se establece un umbral de correlación absoluta τ , en donde una de las variables X_j se selecciona si cumple la siguiente condición:

$$|r_{X_j, has_{fault}}| \geq \tau \quad (5)$$

Para este caso en específico, el umbral τ se define en $\tau = 0.3$, lo que permite excluir variables con baja asociación lineal y de esta forma disminuir la dimensionalidad del problema sin perder información.

3.4.5 Construcción del conjunto de variables del modelo

Las variables que satisfacen el criterio de selección anteriormente mencionado se convierten en un conjunto de variables de entrada X del modelo estadístico de la siguiente forma:

$$\mathbf{X} = [X_1, X_2, \dots, X_k] \quad (6)$$

mientras que la variable *has_fault* se utiliza como variable objetivo del problema de la siguiente manera:

$$\mathbf{y} = \text{has_fault} \quad (7)$$

Con este procedimiento se permite que el modelo este basado de forma exclusiva con variables eléctricas que posean una relación estadísticamente significativa con la ocurrencia de fallos.

3.5 Implementación del Modelo de Predicción de Fallos

Para poder realizar el modelo de predicción de fallos considerando las variables eléctricas seleccionadas a partir del análisis de correlación, se determina un modelo de regresión logística para que se oriente en la clasificación binaria de escenarios con y sin fallos eléctricos. Con este modelo se puede estimar la probabilidad en el que puede generar un fallo mediante el uso de las variables operativas del sistema de distribución.

3.5.1 Selección de variables de entrada

A partir del vector de correlación obtenido en la etapa anterior, se seleccionan aquellas variables cuyo coeficiente de correlación supera el umbral predefinido τ , excluyendo la variable objetivo *has_fault* de la siguiente forma:

$$\mathcal{X}_{sel} = \{X_j \mid r_{X_j, \text{has_fault}} \geq \tau\} \quad (8)$$

Con este criterio se garantiza el uso únicamente de variables con una asociación relevante relacionadas con la presencia de fallos.

3.5.2 Definición de las variables del modelo

Una vez seleccionadas las variables representativas, se define la matriz de entrada \mathbf{X} , el vector objetivo \mathbf{y} a través de la siguiente forma:

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix} \quad (9)$$

$$\mathbf{y} = \begin{bmatrix} \text{has_fault}_1 \\ \text{has_fault}_2 \\ \vdots \\ \text{has_fault}_n \end{bmatrix} \quad (10)$$

donde n representa el número total de escenarios simulados y k el número de variables seleccionadas.

3.5.3 División del conjunto de datos

Con la nueva base de datos seleccionada en base al análisis de correlación, se realiza la separación de datos logrando realizar el entrenamiento y las pruebas del modelo. Para esto, se utiliza un porcentaje de 70 enfocado en el entrenamiento y el 30 se establece en pruebas, lo cual garantizará la proporción de escenarios con y sin fallos mediante una división aleatoria que cumpla la siguiente condición:

$$\mathbf{X} = \mathbf{X}_{train} \cup \mathbf{X}_{test}, \mathbf{y} = \mathbf{y}_{train} \cup \mathbf{y}_{test} \quad (11)$$

3.5.4 Formulación del modelo de regresión logística

Para poder modelar el sistema predictivo propuesto, se realiza la regresión logística la misma que permitirá definir la probabilidad de ocurrencia de un fallo eléctrico como una función sigmoideal de una composición lineal de las variables de entrada, y la cual se expresa de la siguiente forma:

$$P(\text{has_fault}_i = 1 \mid X_i) = \frac{1}{1 + e^{-z_i}} \quad (12)$$

donde:

$$z_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij}$$

Siendo β_0 un término independiente, mientras que β_j representa los coeficientes del modelo asociados a cada variable eléctrica seleccionada.

3.5.5 Entrenamiento del modelo

Para entrenar el modelo se utiliza el conjunto (X_{train}, Y_{train}) , y ajustando los coeficientes β de manera iterativa hasta alcanzar la convergencia. Con este proceso se puede comprender la relación de las variables eléctricas del sistema y la probabilidad en que exista un fallo eléctrico.

3.5.6 Evaluación del modelo

Cuando ya se haya entrenado este proceso se procede a evaluar el modelo a través de un conjunto de datos de prueba (X_{test}, Y_{test}) , en donde se analizará la capacidad para que exista una correcta clasificación y generalización del modelo planteado. El principal resultado de esta propuesta es obtener una probabilidad de fallo que esté asociada con cada escenario y que pueda ser utilizada para identificar las condiciones de ocurrencia de un fallo eléctrico antes de que ocurra.

3.6 Validación del Modelo

Para poder evaluar la capacidad del modelo logístico implementado se utilizan las diferentes métricas de desempeño relacionadas con problemas de clasificación binaria. A continuación, se describe cada una de estas variables.

3.6.1 Matriz de confusión

La matriz de confusión muestra el desempeño cotejando las predicciones obtenidas con los valores reales visualizados y la cual se define de la siguiente forma:

$$\begin{bmatrix} VP & FP \\ FN & VN \end{bmatrix} \quad (13)$$

donde:

- **VP (Verdaderos Positivos):** escenarios con fallo correctamente identificados.

- **FP (Falsos Positivos):** escenarios normales clasificados erróneamente como fallos.
- **FN (Falsos Negativos):** escenarios con fallo no detectados por el modelo.
- **VN (Verdaderos Negativos):** escenarios normales correctamente clasificados.

3.6.2 Exactitud (Accuracy)

Por su parte, la exactitud permite medir la cantidad total de predicciones válidas realizadas por el modelo y se la representa por:

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN} \quad (14)$$

3.6.3 Precisión (Precision)

Permite determinar la confiabilidad para predecir el modelo positivo, eso quiere decir que cuando los escenarios identificados como fallos son correspondientes de manera real a fallos del sistema, siendo la representación:

$$Precision = \frac{VP}{VP + FP} \quad (15)$$

3.6.4 Sensibilidad o Tasa de Verdaderos Positivos (Recall)

Por otro lado, la sensibilidad determina la capacidad que tiene el modelo para poder predecir los diferentes escenarios con fallo eléctrico y se la representa de la siguiente forma:

$$Recall = \frac{VP}{VP + FN} \quad (16)$$

3.6.5 Medida F1 (F1-score)

Por otro lado, la medición de F1 representa el equilibrio entre precisión y sensibilidad, y está definido como una medida armónica entre ambas métricas de la siguiente forma:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (17)$$

3.6.6 Curva ROC y Área Bajo la Curva (AUC)

El parámetro ROC busca evaluar el desempeño que se tiene en el modelo en base a diversos umbrales de decisión, representando como se relaciona la tasa de los positivos verdaderos y los falsos positivos así:

$$TPR = \frac{VP}{VP + FN}, FPR = \frac{FP}{FP + VN} \quad (18)$$

Por su parte, el AUC se conoce como la capacidad discriminativa global del modelo y toma valores entre 0 y 1.

4 Análisis de Resultados

4.1 Variables Eléctricas del Sistema Base

El análisis del sistema IEEE de 33 barras en condiciones normales de operación permitió identificar el comportamiento general del sistema eléctrico planteado previo a la inclusión de los diferentes tipos de fallos. Como se observa en la **Figura 6**, las simulaciones realizadas identificaron un valor promedio de voltaje entre todos los nodos de 0.9481 p.u, lo cual indica que la red opera en rangos aceptable de calidad de servicio. Sin embargo, se aprecian variaciones tanto superiores como inferiores que se consideran propias de la topología radial, principalmente en los nodos más alejados.

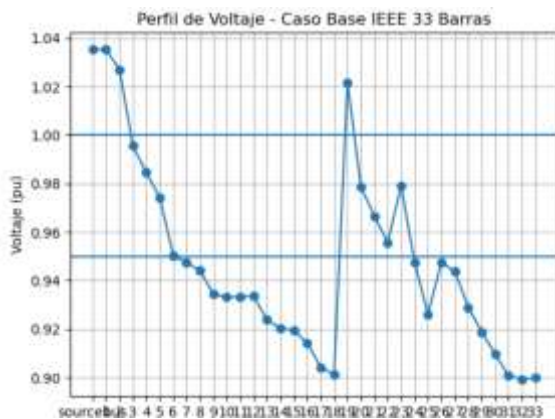


Figura 6. Perfil de Voltaje del Sistema IEEE Simulado en Estado Normal

De igual manera, se observa en la **Tabla 5** un valor de voltaje mínimo de 0.8992 p.u que se ubica en la barra 32, lo cual muestra que exista una zona débil eléctricamente debido a su lejanía a la fuente del alimentador que tiene como consecuencia el incremento de la impedancia por las líneas de transmisión y el consumo de potencia reactiva. Por su parte, el voltaje máximo en el sistema alcanzó un valor de 1.0354 p.u, lo cual está dentro de los rangos recomendados, identificando la no existencia de sobrevoltajes en el escenario base.

Tabla 5. Estadísticas de Voltaje en Condiciones Normales del Sistema IEEE 33 Barras

Estadísticas de Voltaje	Resultados
Promedio	0.9481 pu
Mínimo	0.8992 pu (Bus: 32)
Máximo	1.0354 pu

Por otro lado, se verifica en la **Figura 7** la corriente en las líneas, con un valor promedio de 116.2358 A, mínimo de 4.2385, y máxima de 599.0611 A comprobando las variaciones de este parámetro a lo largo de sistema de distribución.



Figura 7. Corrientes de Línea del Sistema IEEE Simulado en Estado Normal

En cuanto a las pérdidas existentes en el sistema, enfocado en la **Figura 8** se obtuvieron pérdidas activas de 351.47 kW y pérdidas reactivas de 249.50 kVAr, valores que representan adecuadamente un sistema con alta resistencia equivalente y cargas distribuidas a lo largo del alimentador. Estos resultados se conforman como un indicador global del desempeño del sistema en estado nominal que sirve de contraste para las comparaciones con respecto al estado de falla.

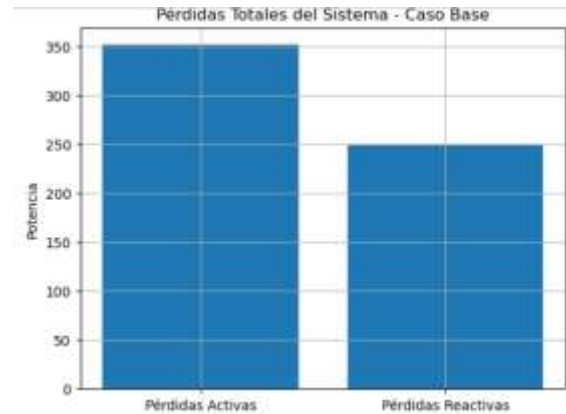


Figura 8. Pérdidas de Potencia del Sistema IEEE Simulado en Estado Normal

Finalmente, se detalla en la **Figura 9** como se realizó las variaciones de la cargabilidad del sistema analizado obteniendo para el 50% un promedio de voltaje de 0.992 p.u; para un 80% de carga 0.965 p.u; un 120% de carga con voltajes de 0.933 p.u; y por último para un 150% de carga con un promedio de 0.911 p.u.

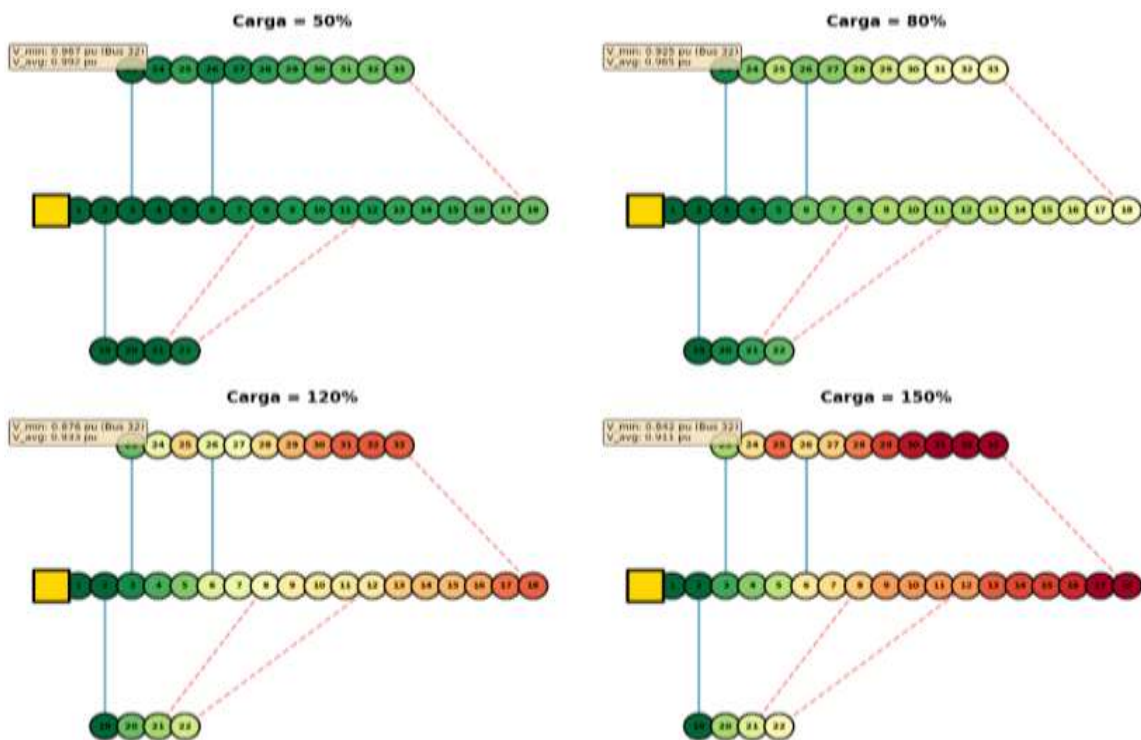


Figura 9. Comparación de Variables Eléctricas con Variación de Cargabilidad del Sistema

4.2 Variables del Sistema con Fallos Eléctricos

El análisis del comportamiento del sistema IEEE de 33 barras bajo condiciones de fallo permitió conocer el impacto de los eventos anómalos aplicados en los nodos del sistema en las variables eléctricas analizadas y de esta forma compararlos con las condiciones de operación normal del sistema.

Se realizó la aplicación de una falla de tipo trifásica en el nodo 8 como se observa en la **Figura 10**, siendo este nodo crítico porque se encuentra en el ramal principal.

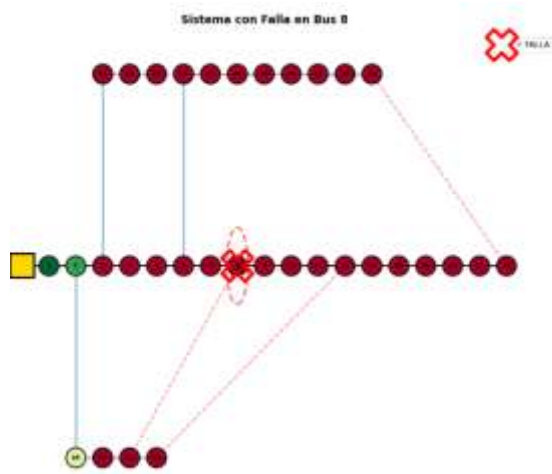


Figura 10. Esquema de Simulación de Falla en el Nodo 8

En este escenario, se observó que existe una caída demasiado elevada de voltaje llegando a un valor cercano a 0.0005 p.u como se observa en la **Tabla 6**, lo cual representa una disminución de aproximadamente 0.8987 p.u con respecto al valor mínimo obtenido durante la operación normal. Este resultado corresponde a que el fallo eléctrico localizado en este nodo se propaga de manera significativa a lo largo de todo el sistema, verificando la sensibilidad del modelo ante fallos ubicados incluso en las barras más alejadas del nodo principal.

Tabla 6. Estadísticas de Voltaje en Condiciones de Falla en el Nodo 8

Estadísticas de Voltaje en Falla 1	Resultados
Promedio	0.8992 p.u
Mínimo	0.0005 p.u
Caída de Voltaje	0.8987 p.u

Por su parte, la inclusión de falla en el nodo 18, se visualiza en la **Figura 11** tiene como resultados que el voltaje mínimo del sistema durante el fallo se mantiene prácticamente sin variaciones con respecto al escenario base, registrando valores cercanos a 0.8992 p.u.

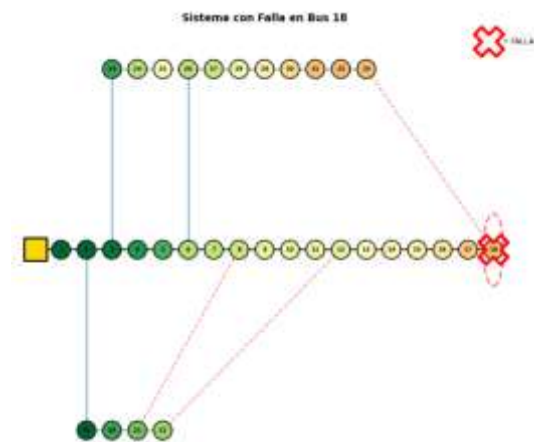


Figura 11. Esquema de Simulación de Falla en el Nodo 18

Dentro de la **Tabla 7** se verifica los resultados estadísticos de voltaje con respecto a la presencia de falla en el nodo 18, el cual mantiene el voltaje promedio con respecto al voltaje mínimo.

Tabla 7. Estadísticas de Voltaje en Condiciones de Falla en el Nodo 18

Estadísticas de Voltaje en Falla 2	Resultados
Promedio	0.8992 p.u
Mínimo	0.8992 p.u
Caída de Voltaje	-0.0000 p.u

A su vez dentro de la **Figura 12** se exhibe la topología del sistema incluyendo la falla dentro de unos de los ramales secundarios que para este caso es el nodo 25.

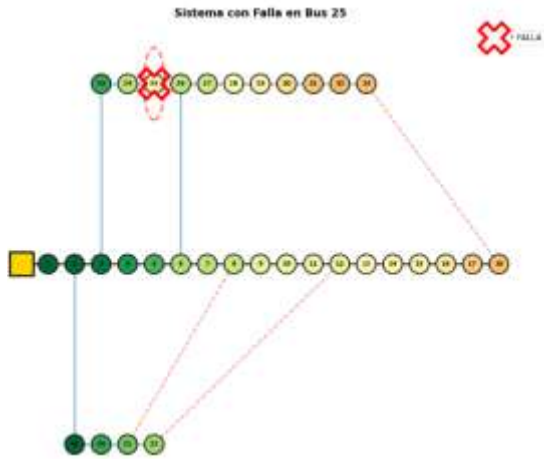


Figura 12. Esquema de Simulación de Falla Monofásica en el Nodo 25

Al igual que los resultados anteriores se define dentro de la **Tabla 8** los resultados estadísticos de voltaje con respecto a la presencia de falla en el nodo 18, en donde se observa un voltaje promedio de 0.8992 p.u que se asemeja al voltaje mínimo

encontrado verificando que en este nodo se presenta una caída de voltaje de 0.00 p.u.

Tabla 8. Esquema de Simulación de Falla en el Nodo 25

Estadísticas de Voltaje en Falla 3	Resultados
Promedio	0.8992 p.u
Mínimo	0.8992 p.u
Caída de Voltaje	0.0000 p.u

Este comportamiento en ambos nodos se da debido a que, por la localización del fallo y la redistribución de corrientes, atenúa el impacto del fallo sobre el resto del sistema.

Una vez comprobado el comportamiento del sistema ante fallos, se procedió a generar los escenarios de prueba totales para la generación de datos eléctricos que permitan aplicar la correlación. La misma se presenta en la **Figura 13**, se generan diversos tipos de fallos tanto monofásicos como trifásicos generando con aleatoriedad las ubicaciones donde se produce la anomalía de la red.

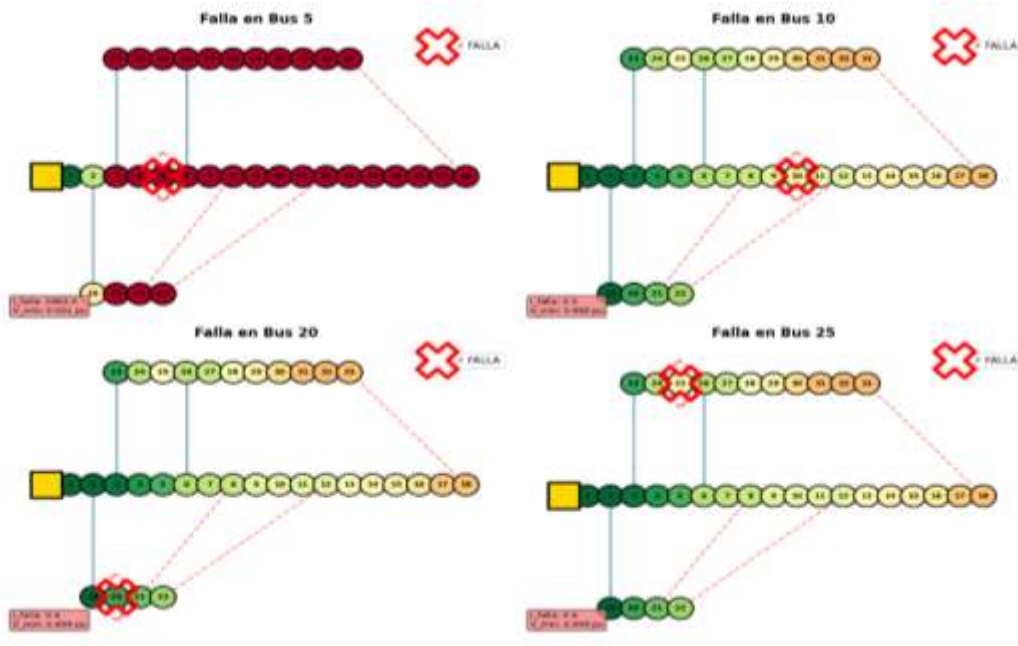


Figura 13. Generación de Múltiples Fallos en Nodos Aleatorios

En base a ello, se obtuvieron los resultados aplicados en la **Tabla 9** dentro de la que se pudo evidenciar el voltaje mínimo promedio en condiciones normales de operación fue de 0.8998 p.u, mientras que en los escenarios donde se simuló una falla se obtuvieron valores promedio de 0.9006 p.u, lo cual genera una diferencia casi despreciable de -0.0008 p.u. Este resultado indica que los fallos provocados dentro del sistema no reflejan reducciones sostenidas del voltaje mínimo, lo cual hace indispensable identificar patrones multivariable para poder modelar las condiciones precursoras de fallos.

Con respecto a las corrientes del sistema, se verificó una corriente máxima promedio de 596.26 A en operación normal, mientras que para los escenarios de fallo se obtuvo una disminución a los 592.59 A, lo cual representa una variación aproximada del -0.6% . Esta disminución se puede asociar a la desconexión de cargas debido a la presencia del fallo lo que genera una redistribución de flujos de corriente. Como en el caso del voltaje, este comportamiento hace evidente la necesidad de involucrar múltiples variables para poder realizar la detección de un fallo por medio de modelos estadísticos.

Tabla 9. Estadísticas Eléctricas del Sistema en Múltiples Escenarios de Fallos

Estadísticas Eléctricas de los Múltiples Escenarios	Resultados
Voltaje Mínimo Promedio en Estado Normal	0.8998 p.u
Voltaje Mínimo Promedio en Estado de Falla	0.9006 p.u
Diferencia de Voltaje	-0.0008 p.u

Corriente Máxima Promedio en Estado Normal	596.26 A
Corriente Máxima Promedio en Estado Con Falla	592.59 A
Incremento de Corriente	-0.6%

Se realizó la distribución de los voltajes mínimos como se presenta en la **Figura 14** en donde se aprecia que la mayor parte de los escenarios simulados poseen valores que se encuentran aproximadamente entre 0.87 y 0.93 p.u, con una mayor concentración en 0.90 p.u. Este comportamiento comprueba que el sistema IEEE de 33 barras opera de forma reiterativa cerca del límite inferior de voltaje, principalmente en barras alejadas de la subestación.

De igual manera, se aprecia que todos los valores obtenidos durante la simulación se ubican por debajo del límite inferior de 0.95 p.u lo cual indica que es sistema, aunque trabaje de manera normal, sin la inclusión de fallos operan en los límites críticos, resaltando la necesidad de identificar condiciones precursoras antes de que se produzcan eventos de fallo.

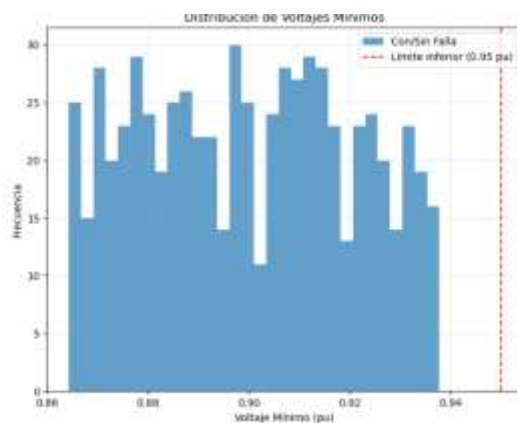


Figura 14. Distribución Estadística de Voltaje Mínimo en los Escenarios Generados

Por otro lado, se comprobó la relación entre corriente máxima del sistema y el mínimo voltaje enfocado en la **Figura 15**, la cual tiene una tendencia lineal negativa bien definida, esto demuestra que a medida que exista un aumento en la corriente el voltaje mínimo disminuye.

Tanto para los escenarios normales como los de presencia de fallo se superponen a lo largo de esta tendencia, lo cual muestra que el incremento de corriente asociado a mayores niveles de carga o a redistribuciones de flujo afecta de manera directa el perfil de voltaje.

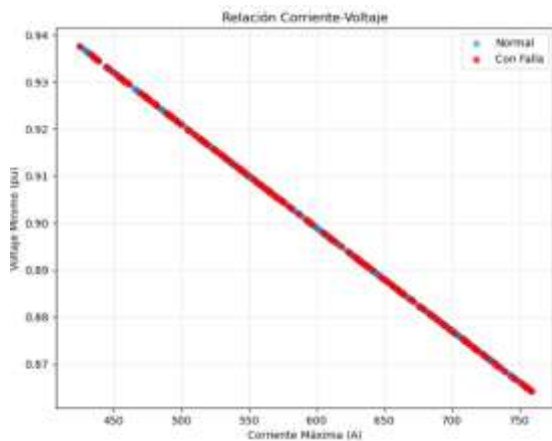


Figura 15. Relación corriente–voltaje mínimo

Finalmente, se comprobó la relación entre las pérdidas técnicas del sistema y la potencia total demandada en la **Figura 16**, en la que se puede evidencia una tendencia creciente no lineal, en donde las pérdidas incrementales de forma acelerada conforme se incrementa la potencia total. Al igual que en la comparación anterior, los escenarios normales y con fallo se distribuyen a lo largo de la misma curva, lo cual muestra que el nivel de pérdidas está fuertemente influenciado por la condición de carga del sistema.

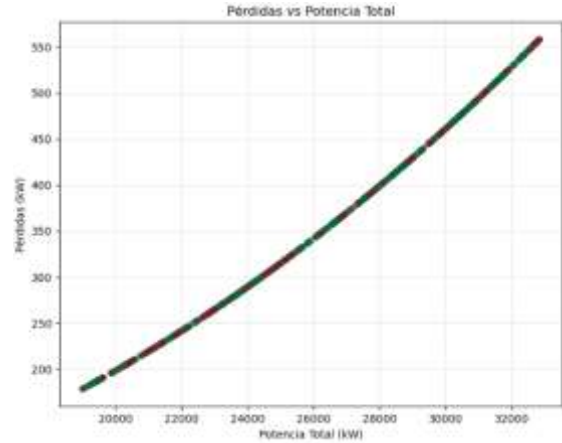


Figura 16. Relación entre pérdidas técnicas y potencia total

4.3 Resultados de Correlación de Pearson

Se procedió a realizar el análisis de correlación el cual tuvo como objetivo reconocer las relaciones presentes entre las variables eléctricas del sistema y la ocurrencia de fallos, la cual se representa con una variable binaria. En los resultados que se obtuvieron se detalla que las correlaciones presentan magnitudes bajas a moderadas, característica acorde al tipo de fenómeno complejo como la generación de los fallos dentro del sistema en el cual suelen interferir diversos factores de manera simultánea y no necesariamente de forma lineal.

En la **Figura 17** se visualiza el diagrama de calor de la matriz de correlación de Pearson en donde se aprecian dos patrones principales, siendo el primero una correlación positiva entre las variables de `min_voltage_pu`, `avg_voltage_pu` y `max_voltage_pu` que varían de manera constante debido a que todas representan medidas de voltaje del sistema, al igual que las variables de `total_current_A` y `max_current_A` también con una correlación positiva fuerte.

Por otro lado, se evidencia que existe una correlación fuerte negativa entre la variable de voltaje y las de corriente, lo cual muestra

el comportamiento físico propio de los sistemas eléctricos.

Las variables `load_mwtt`, `losses_kW` y `has_fault` presentan correlaciones más débiles con el resto de las variables, lo que indican comportamientos más independientes entre sí.

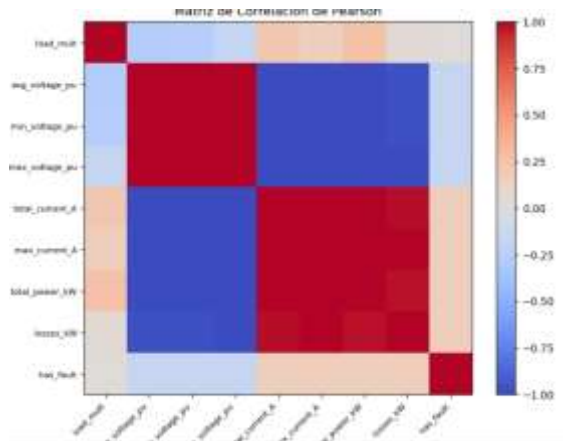


Figura 17. Diagrama de Calor de la Matriz de Correlación de Pearson

En la **Tabla 10**, se evidencia las variables que poseen una mayor correlación positiva con la presencia de fallos son la corriente máxima del sistema, las pérdidas activas, la corriente y la potencia totales demandada, con coeficientes de correlación cercanos a 0.16. Pese a que estos valores no indican una relación fuerte, evidencian su influencia en el estrés eléctrico del sistema, asociados a una mayor probabilidad de ocurrencia de fallos.

Por otro lado, las variables referentes al perfil de voltaje, tales como el voltaje promedio, mínimo y máximo, tienen correlaciones negativas con la ocurrencia de fallos, con valores cercanos a -0.16 . Estos valores demuestran que en niveles menores de voltaje se genera una relación ligera con una mayor probabilidad de fallo, siendo esta una muestra asociada a condiciones de debilidad eléctrica o alta impedancia en el sistema.

Tabla 10. Resultados de Correlación de Pearson más Representativos

Variable	Correlación Obtenida
<code>has_fault</code>	1.000000
<code>max_current_A</code>	0.164628
<code>losses_kW</code>	0.164485
<code>total_current_A</code>	0.163784
<code>total_power_kW</code>	0.162539
<code>load_mult</code>	0.019189
<code>avg_voltage_pu</code>	-0.162385
<code>min_voltage_pu</code>	-0.162816
<code>max_voltage_pu</code>	-0.164560

4.4 Métricas del Modelo de Predicción Planteado

Una vez definidas las variables con mayor relación, se realizó el planteamiento del modelo de regresión logística el cual tiene como objetivo anticipar la ocurrencia de fallos eléctricos. Para ello, se seleccionó datos de manera aleatoria un subconjunto a prueba del 30% del total de los datos, tomando en consideración el desbalance natural entre escenarios normales y escenarios con fallo. Para poder mitigar este efecto, se utilizó un modelo basado en ponderación de clases, priorizando la correcta identificación de eventos de fallo.

Como se observa en la **Tabla 11** se generó un reporte de las pruebas de clasificación en donde se visualiza que el modelo de regresión logística planteado posee un desempeño asimétrico entre las clases analizadas.

Para la clase de falla, el modelo alcanza un recall de 1.00, lo cual indica que todos los escenarios de fallo se identificaron de manera correcta, pero con una precisión baja de 0.23 debido al elevado número de falsos positivos. Por otro lado, la clase de operación normal no comprueba que no es

correctamente reconocida, debido a que el modelo entrega valores nulos con precisión, recall y F1-score. La exactitud global del modelo es del 23%, un valor muy bajo con respecto a otro tipo de sistemas de clasificación general, pero que mantiene la coherencia para la detección de fallos.

Tabla 11. Resultados de Evaluación del Modelo de Predicción

Clase	Precisión	Recall	F1-score	Soprote
Normal (0)	0.00	0.00	0.00	30
Fallo (1)	0.23	1.00	0.38	9
Accuracy	—	—	0.23	39
Macro avg	0.12	0.50	0.19	39
Weighted avg	0.05	0.23	0.09	39

Por su parte, la matriz de confusión definida en la **Tabla 12**, observa que el modelo de regresión logística clasifica de manera correcta a los escenarios evaluados como fallos. De los 39 escenarios seleccionados para la evaluación, los 30 de operación normal fueron correctamente identificados, es decir verdaderos positivos, al igual que los 9 escenarios de fallo fueron correctamente detectados. No se falsos positivos ni falsos negativos dentro de esta evaluación.

Desde la detección de fallos, este comportamiento se puede verificar que la sensibilidad (recall) del 100 %, implica que el modelo no omite ningún evento de fallo. Este resultado se considera importante debido a que en aplicaciones como el mantenimiento preventivo, esta condición permite detectar condiciones potencialmente críticas en el sistema.

Tabla 12. Matriz de Confusión del Sistema de Clasificación

Clase real / Clase predicha	Normal (0)	Fallo (1)
Normal (0)	30	0

Fallo (1)	0	9
-----------	---	---

Por otro lado, se evaluó el modelo mediante la curva ROC que se observa en la **Figura 18**, se observa un rendimiento deficiente del modelo con un AUC de 0.40, que está por debajo del valor de referencia de 0.50. Este valor indica que el modelo en ciertas ocasiones el modelo no genera clasificaciones correctas presentando falsos positivos principalmente en condiciones de operación normal.

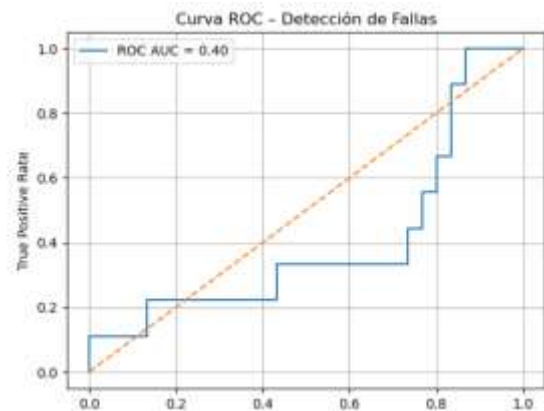


Figura 18. Curva ROC de Detección de Fallos

Por su parte, en la **Figura 19** se visualiza la curva de Precisión del Modelo la cual tiene un valor de 0.29, oscilando entre 0.2 a 0.3 manteniéndose constante durante la mayor parte del rango de recall. Definiendo de esta forma que se dispone de una capacidad limitada para poder nidificar de manera correcta las fallas, generando ciertos falsos positivos cuando intenta detectar condiciones anómalas particulares.

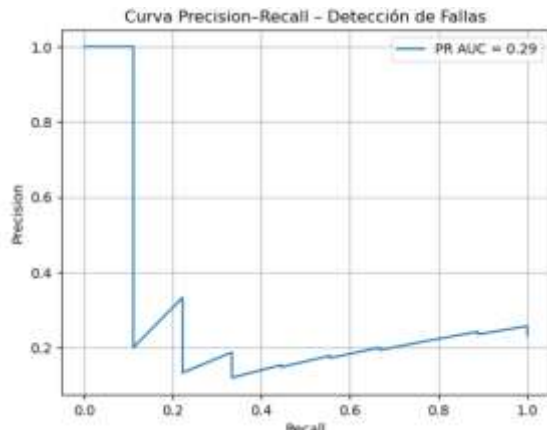


Figura 19. Curva de Precisión del Modelo Planteado

Finalmente, como se realizó la identificación de probabilidades del modelo que ha sido planteado, analizado en la **Figura 20**. Este resultado muestra que existe cierto solapamiento con respecto a las probabilidades predichas para ambas clases. La mayoría de las predicciones, tanto para casos de "No falla" (azul) como de "Falla" (naranja), se concentran en el rango de probabilidades bajas (0.4 a 0.5), lo cual indica que el modelo a veces se puede confundir entre las dos clases.

Existe una pequeña cantidad de casos de falla en donde el modelo predice con alta probabilidad, pero en la gran mayoría de las predicciones se mezclan en rangos similares. Esta escasa separación entre las distribuciones de probabilidad explica el por qué las métricas ROC y Precision-Recall han obtenido valores bajos, requiriendo ampliar la cantidad de datos de entrenamiento para poder mejorar el aprendizaje efectivo entre fallas y operación normal del sistema eléctrico.

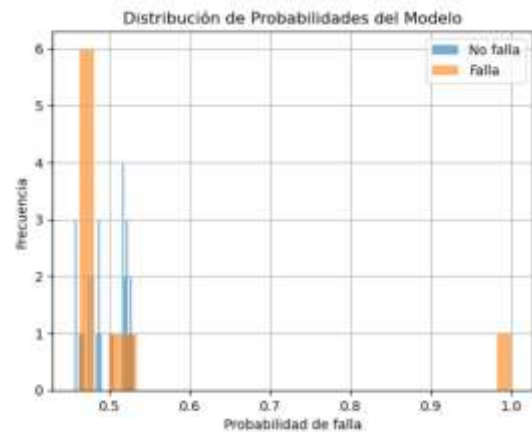


Figura 20. Distribución de Probabilidades del Modelo

5 Discusión de Resultados

Los resultados obtenidos en la simulación del sistema IEEE de 33 barras en condiciones normales de operación muestran un perfil de voltaje promedio de 0.9481 p.u, junto con un valor mínimo de 0.8992 p.u ubicado en la barra 32, lo que confirma la existencia de zonas eléctricamente débiles relacionadas con la distancia a la subestación y al incremento progresivo de la impedancia de las líneas. De igual forma, las pérdidas activas y reactivas obtenidas (351.47 kW y 249.50 kVAr) indican un sistema con cargas distribuidas y alta resistencia equivalente, lo cual valida al modelo seleccionado como base para el análisis estadístico realizado.

Por otro lado, la ejecución de simulaciones bajo distintos escenarios operativos, en donde se incluyeron variaciones de cargabilidad e introducción de fallas monofásicas y trifásicas en diferentes nodos, permitió evaluar la respuesta del sistema frente eventos anómalos. Los resultados demuestran que las fallas ubicadas en nodos críticos del ramal principal, como el nodo 8, producen caídas severas de voltaje cercanas a cero, mostrando un impacto sistémico significativo. De manera contraria, las fallas localizadas en ramales secundarios como el 18 y 25 tienen efectos atenuados sobre las

variables globales del sistema, esto por la redistribución de corrientes y la topología de la red. Esta muestra que la severidad del impacto de una falla no depende únicamente de su tipo, sino también de su localización dentro del sistema.

Por su parte, la aplicación del análisis de correlación de Pearson identificó relaciones de magnitud baja a moderada, entre las variables eléctricas del sistema y la presencia de fallos. Los resultados obtenidos indican correlaciones positivas cercanas a 0.16 entre la presencia de fallos y variables como la corriente máxima, las pérdidas activas y la potencia demandada, lo que indica que condiciones de mayor estrés eléctrico aumentan de forma ligera la probabilidad de eventos de fallo. Por otro lado, las correlaciones negativas encontradas entre los perfiles de voltaje y la variable de fallo indican que niveles menores de voltaje están asociados, a condiciones precursoras de fallas. Estas correlaciones tienen coherencia con la naturaleza multifactorial de los fallos eléctricos, donde ningún parámetro aislado resulta determinante, sino que más bien el fenómeno emerge de la interacción simultánea de múltiples variables operativas.

Finalmente, los resultados del modelo de regresión logística planteado muestran un aspecto importante que se puede relacionar en aplicaciones de mantenimiento preventivo ya que presenta una alta sensibilidad en la detección de fallos. El recall del 100 % para la clase de falla indica que el modelo tiene la capacidad de identificar todos los eventos de fallo simulados, evitando el no detectar situaciones críticas que puedan derivar en interrupciones no planificadas del servicio. Pese a esto, la baja precisión y el bajo AUC de la curva ROC indican una elevada tasa de falsos positivos y una limitada capacidad de discriminación entre estados normales y de fallo. Aunque se han presentado ciertas

limitaciones, los resultados indican que la metodología propuesta puede aplicarse como herramienta complementaria en apoyo a la toma de decisiones, destinada a la identificación temprana de condiciones de riesgo.

6 Conclusiones

El análisis del sistema IEEE de 33 barras en condiciones normales de operación permitió caracterizar de forma adecuada el comportamiento eléctrico en condiciones operativas nominales. Los resultados obtenidos muestran que se tiene un voltaje promedio del sistema de 0.9481 p.u., valor aceptable para este tipo de modelos. Sin embargo, se identificó un valor de 0.8992 p.u. para el voltaje mínimo en la barra 32, lo cual evidencia que existen zonas eléctricamente débiles relacionadas con la topología radial y la lejanía del nodo con respecto a la subestación. Con respecto a las pérdidas activas y reactivas, se obtuvieron valores de 351.47 kW y 249.50 kVAr, respectivamente, comprobando que el sistema posee una alta resistencia equivalente.

Por otro lado, la simulación de los escenarios con variaciones operativas y diferentes tipos de fallas permitió evaluar la respuesta del sistema ante eventos anómalos. Se analizaron condiciones de cargabilidad desde el 50 % hasta el 150 %, observándose una reducción progresiva del voltaje promedio desde 0.992 p.u. en el caso de carga al 50 %, hasta 0.911 p.u. para el caso de carga al 150 %, mostrando un incremento del estrés eléctrico bajo cargas elevadas. Con respecto a los fallos eléctricos provocados, la falla trifásica ubicada en el nodo 8 generó una caída extrema del voltaje mínimo hasta 0.0005 p.u., que representa una disminución de 0.8987 p.u. en relación con el escenario base, comprobando que existe una alta sensibilidad del sistema ante fallos en nodos críticos del ramal principal.

Por su parte, la aplicación del análisis de correlación de Pearson permitió identificar las variables eléctricas con mayor relación estadística con la presencia de fallos eléctricos. Los resultados hallados se pudo evidencia que las correlaciones obtenidas han sido de baja a moderada magnitud, siendo las variables con mayor correlación positiva la corriente máxima con un índice de 0.1646, las pérdidas activas con 0.1645, la corriente total con 0.1638 y la potencia total demandada con 0.1625. Por otro lado, las variables asociadas al perfil de voltaje presentaron correlaciones negativas cercanas a -0.16 , lo cual indica que para menores niveles de voltaje se asocian a una mayor probabilidad de fallo.

Finalmente, se realizó la implementación del modelo de predicción, cuyos resultados demostraron que la metodología propuesta es factible como herramienta complementaria para la planificación del mantenimiento preventivo en sistemas de distribución reales. Para el modelo de la regresión que fue implementado se evidencio que alcanzó un recall del 100% en el fallo eso quiere decir que fueron detectados a tiempo y de manera correcta. La curva ROC con un AUC de 0.40 y la superposición de probabilidades en las predicciones demostraron que el modelo tiene una limitada capacidad de diferenciación entre ciertas condiciones operativas debido principalmente a la baja separación estadística de los datos entre estados normales y de fallo. Pese a esto, la funcionalidad de no omitir fallos por parte del modelo planteado resulta eficiente en aplicaciones de mantenimiento preventivo, lo cual confirma que el modelo propuesto sirve como herramienta complementaria para la toma de decisiones operativas en sistemas eléctricos de distribución.

7 Trabajos Futuros

Se propone ampliar la metodología para poder incluir modelos automáticos de

aprendizaje a través de redes neuronales las cuales deberán tener capacidad para capturar las relaciones no lineales entre la variable de ocurrencia de fallos y la eléctrica. De igual forma, se recomienda integrar variables adicionales relacionadas con la dinámica del sistema, tales como tasas de variación de voltaje y corriente, armónicos y eventos transitorios, para mejorar el conjunto de datos y actualizar el modelo. Finalmente, se plantea integrar el modelo estadístico con sistemas de supervisión en tiempo real los cuales permitan implementar una herramienta inteligente de apoyo en la toma de decisiones para el mantenimiento preventivo y predictivo dentro de un sistema de distribución real.

8 Referencias

- [1] A. H. M. Ibrahim, S. K. Sadanandan, T. Ghaoud, V. S. Rajkumar y M. Sharma, «Incipient Fault Detection in Power Distribution Networks: Review, Analysis, Challenges, and Future Directions,» *IEEE Access*, vol. 12, p. 112822–112838, 2024.
- [2] G. Tahiliani y A. R. Gupta, «Electrical Distribution System Analysis with Atom Search Optimization based DG and DSTATCOM Allocation,» *2022 IEEE Delhi Section Conference (DELCON)*, pp. 1-6, 2022.
- [3] A. Shaheen, R. El-Sehiemy, S. Kamel y A. Selim, «Optimal Operational Reliability and Reconfiguration of Electrical Distribution Network Based on Jellyfish Search Algorithm,» *Energies*, vol. 15, n° 6994, 2022.
- [4] Q. L. e. al., «Incipient Fault Detection in Power Distribution System: A Time-Frequency Embedded Deep

- Learning Based Approach,» *arXiv*, 2023.
- [5] H. Rehman, M. Tariq, A. Sarwar, W. Alhosaini, M. Hossain y S. Batiyah, «Single-Phase Fault Tolerant Multilevel Inverter Topologies—Comprehensive Review and Novel Comparative Factors,» *Energies*, vol. 15, n° 9319, 2022.
- [6] J. S. e. al., «A fault diagnosis and location method for power grid simulators based on voltage threshold and MCNN,» *EAI Endorsed Transactions on Energy Web*, vol. 12, 2025.
- [7] B. Su, Y. Xi, S. Li y B. Yuan, «Research on Optimal Control Strategies on Distribution Network Power Transfer Under Extreme Weather Conditions,» *Electronics*, vol. 14, n° 19, p. 3854, 2025.
- [8] L. Jin, Y. Mao, X. Wang, P. Shi, L. Lu y Z. Wang, «Optimization-Based Maximum-Torque Fault-Tolerant Control of Dual Three-Phase PMSM Drives Under Open-Phase Fault,» *IEEE Transactions on Power Electronics*, vol. 38, n° 3, pp. 3653-3663, 2023.
- [9] D. Anggara, D. Danial y Z. Abidin, «Analysis of the Effect of Increasing Load on Voltage Profile in Electric Power Distribution System,» *Action Research Literate*, vol. 9, n° 9, p. 2095–2103, 2025.
- [10] Y. X. H. H. W. e. a. Chen, «A Large-Scale Dataset of Distributed Renewable Energy Scenarios on the IEEE-33 Bus Network,» *Sci Data*, 2025.
- [11] S. Liu, Z. Tang, B. Chai y Z. Zeng, «Robust Distribution System State Estimation with Physics-Constrained Heterogeneous Graph Embedding and Cross-Modal Attention,» *Processes*, vol. 13, n° 10, p. 3073, 2025.
- [12] A. V. Ramana, U. Bhoga, R. K. Dhulipalla, A. Kiran, B. D. Chary y P. C. S. Reddy, «Abnormal Behavior Prediction in Elderly Persons Using Deep Learning,» *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)*, pp. 1-5, 2023.
- [13] V. K. Hariharan, A. Geetha, F. Granelli y M. G. Nair, «Machine Learning Techniques for Fault Detection in Smart Distribution Grids,» *Energies*, vol. 18, n° 19, p. 5179, 2025.
- [14] V. Rizeakos, A. Bachoumis, N. Andriopoulos, M. Birbas y A. Birbas, «Deep learning-based application for fault location identification and type classification in active distribution grids,» *Applied Energy*, vol. 338, n° 120932, 2023.
- [15] E. V. Zhillin y D. D. Levin, «Analysis of the effectiveness of relative methods for power cable fault detection,» *iPolytech Journal*, vol. 29, n° 3, p. 363–375, 2025.
- [16] W. Hassen, «Diagnosis and Monitoring of Electrical Wiring and Interconnection Systems,» *Aircraft Manufacturing, Safety and Control*, 2025.
- [17] J. Li y D. Gu, «Intelligent Building Electrical Fault Diagnosis Based on Relevant Vector Machine Optimized by Intelligent Algorithm,» *IEEE Transactions on Electrical and Electronic Engineering*, 2025.

- [18] Z. Li, Y. Yang, Y. Yang y D. Wang, «A weighted Pearson correlation coefficient based multi-fault comprehensive diagnosis for battery circuits,» *Journal of Energy Storage*, vol. 60, n° 106584, 2023.
- [19] M. Bartolomeo, P. Varilone y P. Verde, «Probabilistic Estimation of During-Fault Voltages of Unbalanced Active Distribution: Methods and Tools,» *Energies*, vol. 18, n° 18, p. 4791, 2025.
- [20] L. Kong y H. Nian, «Fault Detection and Location Method for Mesh-Type DC Microgrid Using Pearson Correlation Coefficient,» *IEEE Transactions on Power Delivery*, vol. 36, n° 3, p. 1428–1439, 2021.
- [21] D. S. Lestari, E. D. Nursita y R. F. Wijanarko, «Optimization Analysis of Protection Coordination in Loop Distribution Systems with Integrated Distributed Generation Using the Firefly Algorithm and Conventional Methods,» *Jurnal Ilmiah Telsinas Elektro, Sipil dan Teknik Informasi*, vol. 8, n° 2, p. 115–126, 2025.
- [22] L. Becerra, «Estrategia de detección e identificación de fallas eléctricas para sistemas de distribución eléctrica con generación distribuida basado en redes neuronales artificiales,» *Repositorio UPS*, vol. [Tesis de Grado], 2023.
- [23] C. Iberoamericano, «Sistemas De Distribución De Energía Eléctrica,» 18 12 2025. [En línea].
- [24] R. Gomez, D. Cabrera y P. Robles, «Estudio para la Localización de Fallas en Sistemas de Distribución Eléctrica,» *Ingenius. Revista de Ciencia y Tecnología*, n° 30, 2023.
- [25] M. Saltos, A. Velásquez, M. Aguirre, D. Ortíz, A. Villamarín y R. Haro, «Planificación Óptima de Recursos Energéticos Distribuidos para Mejorar la Resiliencia de Sistemas de Distribución de Energía Eléctrica frente a Desastres Naturales: Caso en Lahares Volcánicos,» *Revista Técnica energía*, vol. 18, n° 2, 2022.
- [26] G. Huera, «Fallas en sistemas eléctricos,» 21 01 2020. [En línea]. Available: <https://siticed.com.mx/fallas-en-sistemas-electricos/>.
- [27] C. Moreno, «Detección de eventos en variables eléctricas de sistemas eléctricos de potencia,» *Repositorio UNAL*, vol. [Trabajo de Maestría], 2024.
- [28] V. Reyes, «Análisis de riesgo de fallas en cadenas de aisladores por descargas indirectas,» *Repositorio UNIANDDES*, vol. [Tesis de Grado], 2022.
- [29] D. Flores, «Identificación de causa raíz de fallas por descargas eléctricas en líneas de transmisión,» *Repositorio UPS*, vol. [Tesis de Grado], 2021.
- [30] M. Pérez, S. Rincón, R. Jiménez, J. González y E. Villalvazo, «Desarrollo de una herramienta de análisis estadístico aplicado al pronóstico de la demanda en sistemas eléctricos de potencia,» *Revista Científica de la Facultad de Ingeniería y Ciencias Agropecuarias de la Universidad Nacional de San Luis*, n° 2, pp. 74 - 101, 2025.
- [31] M. Rincón, «Modelo estadístico para la predicción de fallas potenciales

- en redes eléctricas de distribución,» *Repositorio Universidad Tecnológica de Pereira*, vol. [Tesis de Grado], 2019.
- [32] B. Odar, «Análisis y detección de fallas en motores eléctricos aplicando algoritmos de inteligencia artificial,» *Repositorio Universidad de Piura*, vol. [Tesis de Grado], 2023.
- [33] R. Llauce, M. Saldaña y M. Olivera, «Manual correlación y regresión lineal,» *Repositorio UNJ*, 2025.
- [34] F. Villacrés, M. Chamba, A. Torres y C. Lozada, «Estrategia Adaptativa para el Alivio de Carga en Sistemas Eléctricos de Potencia Basada en Regresión Lineal,» *Revista Técnica energía*, vol. 22, nº 1, 2025.
- [35] P. Meera y S. Hemamalini, «Optimal Siting of Distributed Generators in a Distribution Network using Artificial Immune System,» *International Journal of Electrical and Computer Engineering*, vol. 7, nº 2, p. 641~649, 2017.

8.1 Matriz de Estado del Arte

Tabla 13: Matriz de estado del arte.

DESARROLLO DE UN MODELO ESTADÍSTICO BASADO EN LA CORRELACIÓN DE PEARSON PARA ANTICIPAR CONDICIONES PRECURSORAS DE FALLOS ELÉCTRICOS EN UN SISTEMA DE DISTRIBUCIÓN DE PRUEBA REPRESENTATIVO																						
	Año	TÍTULO DEL ARTÍCULO	CITAS	TEMÁTICA					FORMULACIÓN DEL PROBLEMA				RESTRICCIONES DEL PROBLEMA			PROPUESTAS PARA RESOLVER EL PROBLEMA				SOLUCIÓN PROPUESTA		
				Detección de fallas incipientes	Análisis de variables eléctricas	Análisis de sistemas de distribución	Señales eléctricas no estacionarias	Diagnóstico estadístico	Localización precisa de fallas	Complejidad de redes modernas	Diagnóstico simultáneo de múltiples fallas	Mejora de la confiabilidad del sistema	Optimización de redes eléctricas	Variabilidad de carga	Alto costo computacional	Dependencia de mediciones	Uso del coeficiente de Pearson	Análisis multivariable	Integración de machine learning	Identificación de variables críticas	Identificación de patrones precursores	Mejora del perfil de voltaje
1	2024	Incipient Fault Detection in Power Distribution Networks: Review, Analysis, Challenges, and Future Directions	3	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
2	2022	Electrical Distribution System Analysis with Atom Search Optimization based DG and DSTATCOM Allocation	2			<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>									<input checked="" type="checkbox"/>		
3	2022	Optimal Operational Reliability and Reconfiguration of Electrical Distribution Network Based on Jellyfish Search Algorithm	10	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>
4	2023	Incipient Fault Detection in Power Distribution System: A Time-Frequency Embedded Deep Learning Based Approach	8	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
5	2022	Single-Phase Fault Tolerant Multilevel Inverter Topologies—Comprehensive Review and Novel Comparative Factors	15	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>
6	2025	A fault diagnosis and location method for power grid simulators based on voltage threshold and MCNN	3	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>

7	2025	Research on Optimal Control Strategies on Distribution Network Power Transfer Under Extreme Weather Conditions	1		☒					☒	☒			☒				☒		☒			
8	2023	Optimization-Based Maximum-Torque Fault-Tolerant Control of Dual Three-Phase PMSM Drives Under Open-Phase Fault	5	☒				☒	☒		☒			☒		☒							
9	2025	Analysis of the Effect of Increasing Load on Voltage Profile in Electric Power Distribution System	4			☒			☒		☒			☒					☒			☒	
10	2025	A Large-Scale Dataset of Distributed Renewable Energy Scenarios on the IEEE-33 Bus Network	4		☒			☒	☒		☒							☒					
11	2025	Robust Distribution System State Estimation with Physics-Constrained Heterogeneous Graph Embedding and Cross-Modal Attention	12		☒				☒	☒		☒	☒									☒	
12	2023	Abnormal Behavior Prediction in Elderly Persons Using Deep Learning	5	☒				☒			☒						☒		☒			☒	
13	2025	Machine Learning Techniques for Fault Detection in Smart Distribution Grids	20			☒		☒						☒									
14	2023	Deep learning-based application for fault location identification and type classification in active distribution grids	16			☒			☒	☒		☒		☒							☒		
15	2025	Analysis of the effectiveness of relative methods for power cable fault detection	3	☒		☒				☒		☒			☒	☒	☒					☒	
16	2025	Diagnosis and Monitoring of Electrical Wiring and Interconnection Systems	2				☒			☒		☒			☒							☒	
17	2025	Intelligent Building Electrical Fault Diagnosis Based on Relevant Vector Machine Optimized by Intelligent Algorithm	1	☒			☒		☒	☒					☒				☒			☒	
18	2023	A weighted Pearson correlation coefficient based multi-fault comprehensive diagnosis for battery circuits	1		☒	☒	☒	☒		☒													☒
19	2025	Probabilistic Estimation of During-Fault Voltages of Unbalanced Active Distribution: Methods and Tools	2	☒					☒				☒						☒				☒
20	2021	Fault Detection and Location Method for Mesh-Type DC Microgrid Using Pearson Correlation Coefficient	4		☒	☒			☒					☒									
21	2025	Optimization Analysis of Protection Coordination in Loop Distribution Systems with Integrated Distributed Generation Using the Firefly Algorithm and Conventional Methods	8		☒			☒				☒			☒				☒				
22	2023	Estrategia de detección e identificación de fallas eléctricas para sistemas de distribución eléctrica con generación distribuida basado en redes neuronales artificiales	7			☒	☒			☒				☒				☒					

23	2025	Sistemas De Distribución De Energía Eléctrica	9					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>		
24	2023	Estudio para la Localización de Fallas en Sistemas de Distribución Eléctrica	7		<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>									<input checked="" type="checkbox"/>		
25	2022	Planificación Óptima de Recursos Energéticos Distribuidos para Mejorar la Resiliencia de Sistemas de Distribución de Energía Eléctrica frente a Desastres Naturales: Caso en Lahares Volcánicos	2	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>							
26	2020	Fallas en sistemas eléctricos	3			<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>								<input checked="" type="checkbox"/>		
27	2024	Detección de eventos en variables eléctricas de sistemas eléctricos de potencia	2		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>					
28	2022	Análisis de riesgo de fallas en cadenas de aisladores por descargas indirectas	15		<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>								
29	2021	Identificación de causa raíz de fallas por descargas eléctricas en líneas de transmisión	5	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>			
30	2025	Desarrollo de una herramienta de análisis estadístico aplicado al pronóstico de la demanda en sistemas eléctricos de potencia	16					<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>											<input checked="" type="checkbox"/>		
31	2019	Modelo estadístico para la predicción de fallas potenciales en redes eléctricas de distribución	4	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>							
32	2023	Análisis y detección de fallas en motores eléctricos aplicando algoritmos de inteligencia artificial	8		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>					<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		
33	2025	Manual correlación y regresión lineal	10	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>										
34	2025	Estrategia Adaptativa para el Alivio de Carga en Sistemas Eléctricos de Potencia Basada en Regresión Lineal	21	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>				
35	2017	Optimal Siting of Distributed Generators in a Distribution Network using Artificial Immune System	2	<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		
CANTIDAD:				16	13	12	6	14	18	13	11	12	10	14	7	9	7	10	9	10	5	8	13

8.2 Resumen de Indicadores

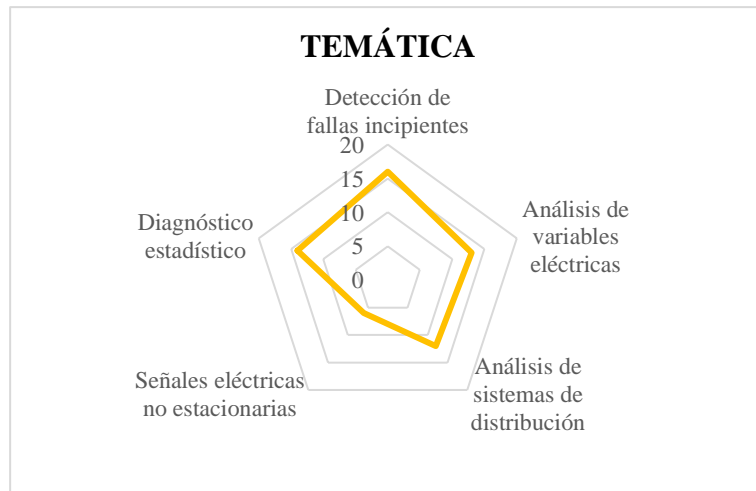


Figura 21. Resumen e indicador de la temática – Estado del arte.

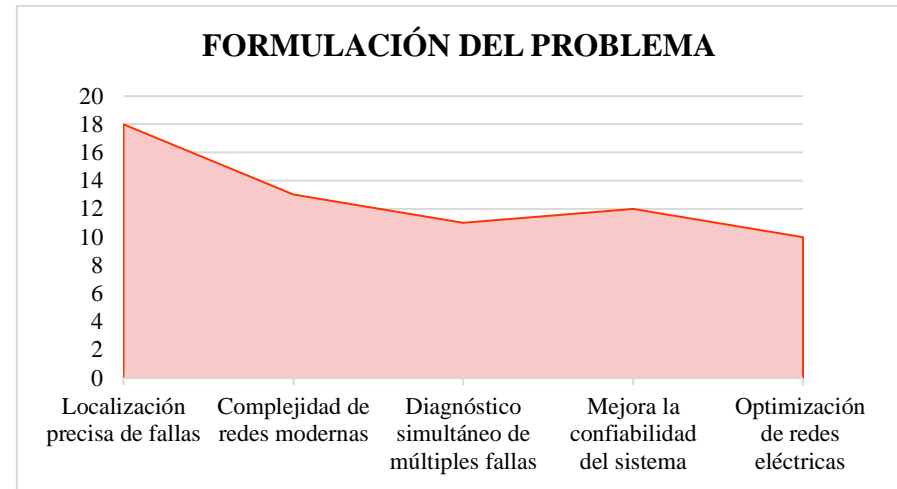


Figura 22. Indicador de formulación del problema - Estado del arte.

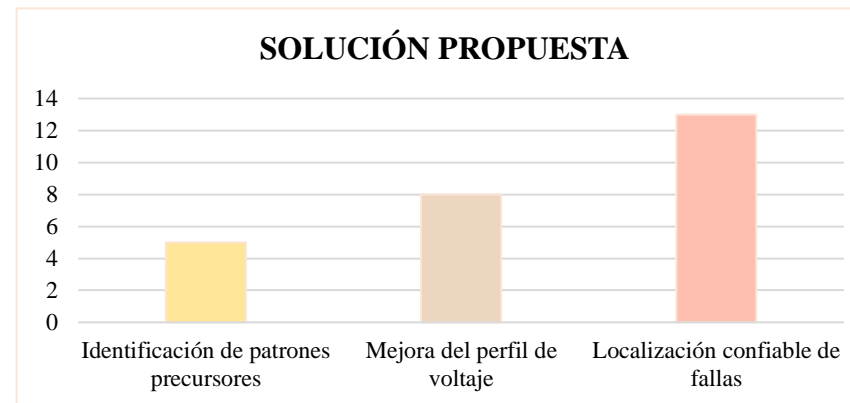


Figura 23. Indicadores de Solución Propuesta – Estado del Arte