

Clasificación de Fallas en Sistemas Eléctricos de Potencia

David Urbaez León, Juan Pablo Botero Sepúlveda, Angela Sofía Rosas, Andres Osorio Santamaria, Manuela Granda Muñoz

Planteamiento del problema

Se busca proponer una solución mediante el análisis de datos y la analítica predictiva que permita facilitar el proceso de detección de fallas eléctricas, clasificando entre fallas internas y fallas externas de los relés de protección para un sistema eléctrico de potencia.

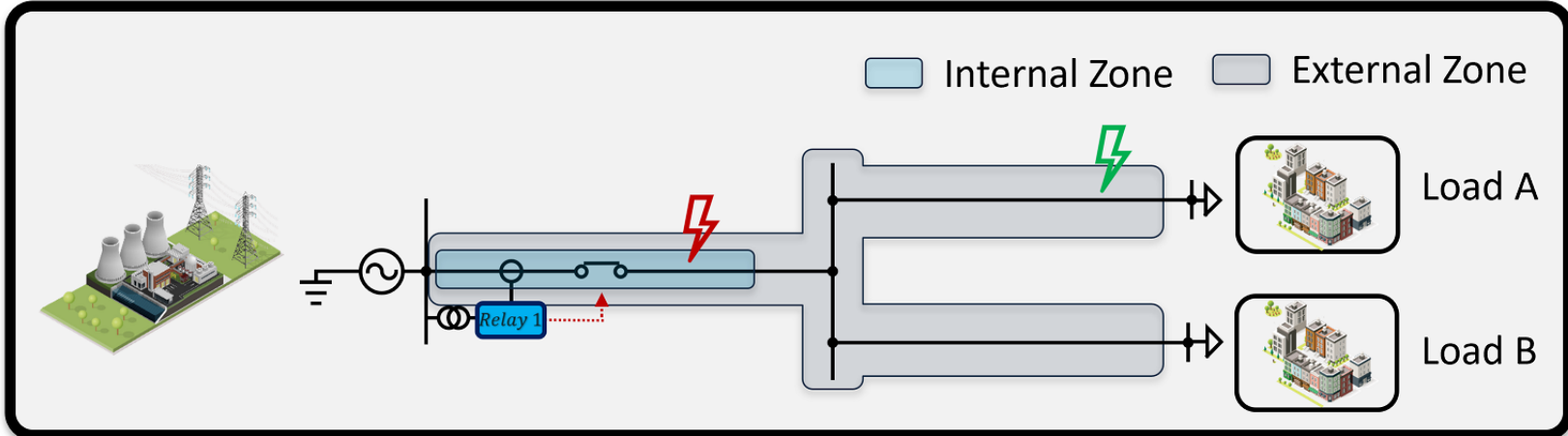


Fig 1.Red para ejemplificar la operación ideal de un relé.

Dataset

El dataset utilizado para el entrenamiento se compone de observaciones de fallas de un sistema eléctrico de 39 barras, que contienen las siguientes características:

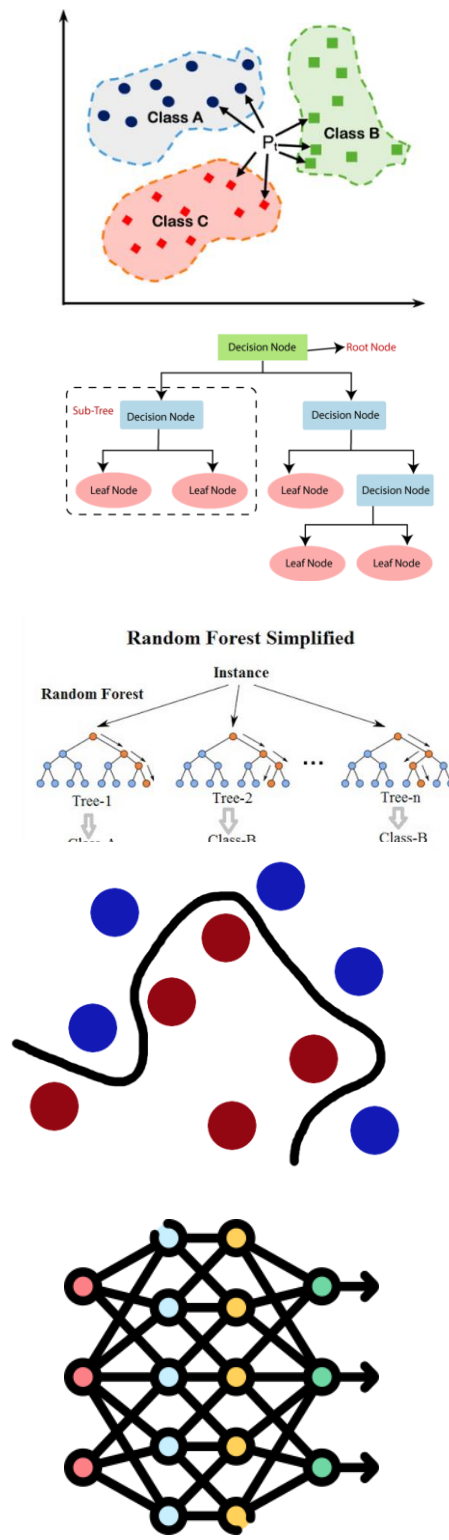
- R: Valor real de la impedancia de falla
- X: Valor imaginario de la impedancia de falla
- V: Voltaje medido en el momento de la falla
- I: Corriente medida en el momento de falla.

La variable objetivo **“inside”** indica si la falla ocurrió en la zona principal de protección de relé, ya que es la señal que indica si se debe o no desconectar la carga eléctrica del sistema, lo que podría significar dejar sin electricidad una parte de un país.

Modelos

Para este proyecto se utilizaran 5 modelos de clasificación diferentes:

- *K Nearest Neighbors*
- *Decision Tree*
- *Random Forest*
- *Support Vector Machine*
- *Red Neuronal*



Además, se realizó un proceso de ajuste de hiper parámetros tales como numero de neuronas, capas, tasa de aprendizaje para el optimizador, entre otras.

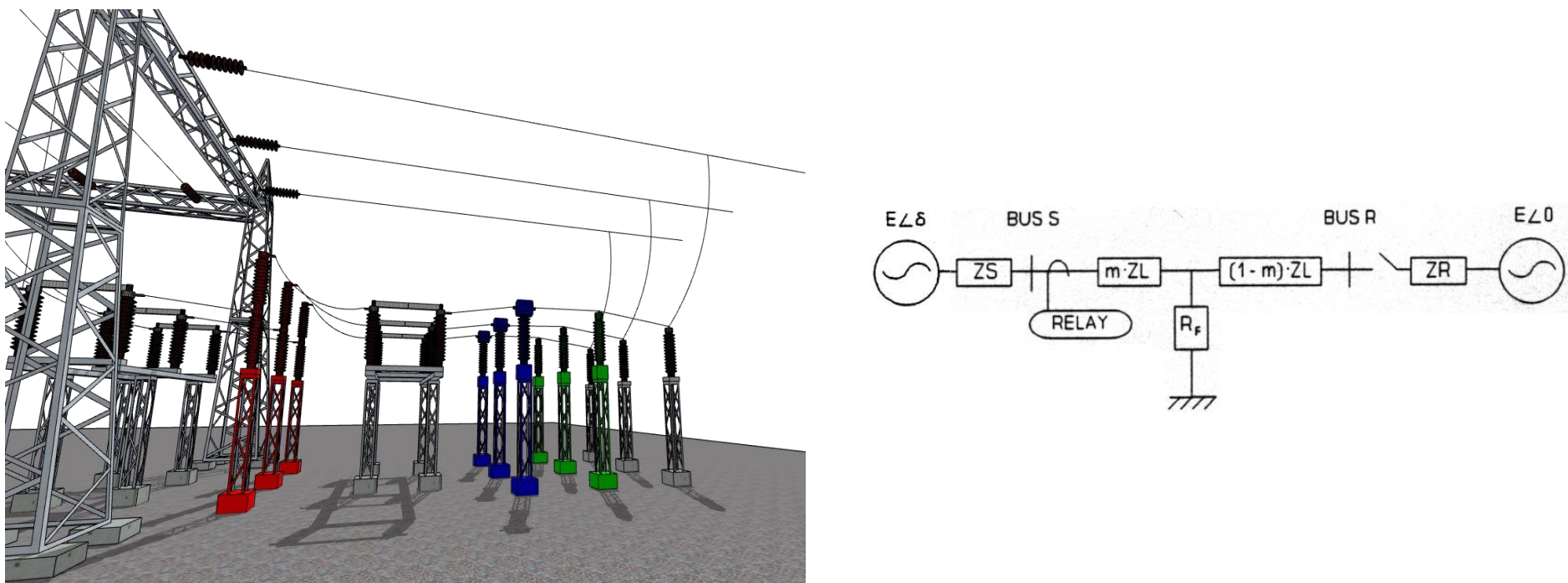


Fig 2.Diagrama para calculo de impedancias de falla

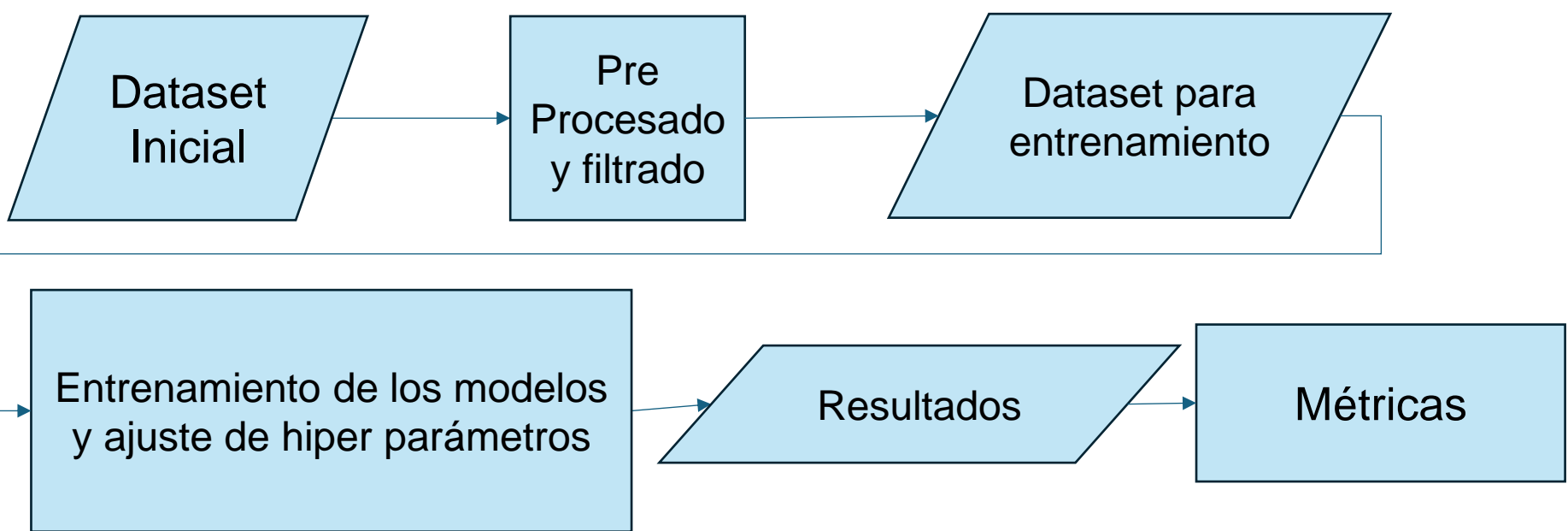


Fig 3.Diagrama del flujo de trabajo.

Resultados

A continuación se muestran las métricas obtenidas con los distintos modelos evaluados.

Métricas				
Method	Accuracy	Recall	Precision	Score F1
KNN	99.76%	0.97	0.98	0.98
Decision Tree	99.94%	0.99	1.00	1.00
Random Forest	100.00%	1.00	1.00	1.00
SVM	99.03%	0.84	1.00	0.91
NN	98.61%	1.00	0.81	0.89

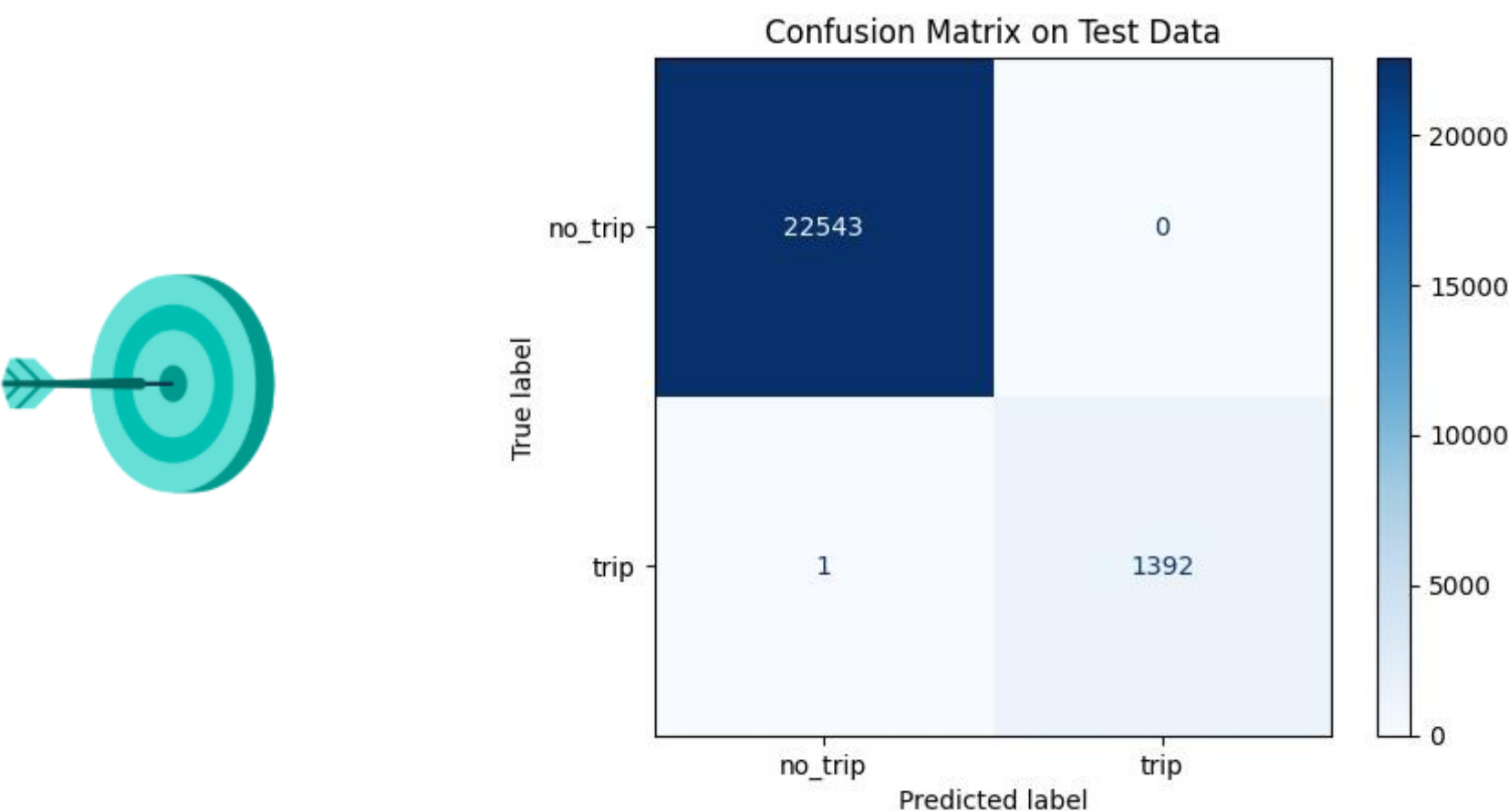


Fig 4. Matriz de confusión sobre datos de test en el modelo random Forest.

Conclusión

El modelo Random Forest obtuvo los mejores resultados entre los modelos considerados, adicionalmente el uso de herramientas de búsqueda de hiperparámetros óptimos como el GridSearchCV o keras tuner permite mejorar significativamente las métricas de desempeño

Sin embargo, se debe tener en cuenta que el dataset tiene muchos mas datos con clasificación negativa, lo cual puede ocasionar un sesgo o un sobreajuste hacia este resultado.

De las métricas se observa que los modelos tuvieron un buen rendimiento general, sin embargo, el tiempo de entrenamiento y los recursos necesarios para realizarlo varían desde unos minutos a un par de horas, por lo que se debe analizar si un modelo mas complejo aporta o no a una mejora de los resultados, en especial cuando se cuenta con recursos de computo limitados.

Referencias

[1] Power System Relaying and Control Committee Subcommittee C – System Protection Working Group C43. Practical Applications of Artificial Intelligence and Machine Learning in Power System Protection and Control

[2] Hanyu Yang, Xubin Liu, Di Zhang, Tao Chen, Canbing Li, Wentao Huang, Machine learning for power system protection and control, The Electricity Journal, Volume 34, Issue 1, 2021,https://doi.org/10.1016/j.tej.2020.106881.