IDENTIFICACION DE FALLAS EN SISTEMAS ELÉCTRICOS DE POTENCIA BASADO EN EL RECONOCIMIENTO DE PATRONES

UNIVERSIDAD POLITÉCNICA SALESIANA SEDE QUITO

CARRERA: INGENIERÍA ELÉCTRICA

Trabajo de titulación previo a la obtención del título de INGENIERO ELÉCTRICO

TEMA: IDENTIFICACION DE FALLAS EN SISTEMAS ELÉCTRICOS DE POTENCIA BASADO EN RECONOCIMIENTO DE PATRONES.

AUTOR: JONATHAN FERNANDO ARAUZ GALLEGOS

DIRECTOR: MILTON GONZALO RUIZ MALDONADO

Quito, febrero 2020

Jonathan Fernando Arauz Gallegos

IDENTIFICACIÓN DE FALLAS EN SISTEMAS ELÉCTRICOS DE POTENCIA BASADO EN EL RECONOCIMIENTO DE PATRONES.

Universidad Politécnica Salesiana Ingeniería Eléctrica

Breve reseña historia e información de contacto:



Jonathan Fernando Arauz Gallegos(Y'1900-M'01). Realizó sus estudios de secundaria en el Colegio Técnico Salesiano 'Don Bosco' "Kennedy" de la ciudad de Quito, se graduó de Bachiller Técnico Industrial en Electricidad y Electrónica. Egresado de la Carrera de Ingeniería Eléctrica de la Universidad Politécnica Salesiana. jarauz@est.ups.edu.ec

Dirigido por:



Milton Gonzalo Ruiz Maldonado (Y'1987-M'01). Milton Ruiz se graduó de la Universidad Politécnica Salesiana en 2012 como Ingeniero Electrónico y en 2015 como Ingeniero Eléctrico; Recibió su M.Sc. en Redes de Comunicaciones en la Pontificia Universidad católica del Ecuador; Sus intereses de investigación están en comunicaciones verdes, infraestructura de medición avanzada, red inteligente, redes inalámbricas, gestión de energía. Es profesor de la Universidad Politécnica Salesiana, Ecuador y parte del Grupo de Investigación GIREI.

mruizm@ups.edu.ec

Todos los derechos reservados:

Queda prohibida, salvo excepción prevista en la ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra para fines comerciales, sin contar con la autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual. Se permite la libre difusión de este texto con fines académicos o investigativos por cualquier medio, con la debida notificación a los autores.

DERECHOS RESERVADOS

©2020 Universidad Politécnica Salesiana

QUITO-ECUADOR

DECLARATORIA DE COAUTORÍA DEL DOCENTE TUTOR/A

Yo, Milton Ruiz Maldonado declaro que bajo mi dirección y asesoría fue desarrollado el trabajo de titulación "*Identificación de Fallas en Sistemas Eléctricos de Potencia Basado en el Reconocimiento de Patrones*" realizado por Jonathan Fernando Arauz Gallegos, obteniendo un producto que cumple con todos los requisitos estipulados por la Universidad Politécnica Salesiana para ser considerados como trabajo final de titulación.

Quito, febrero 2020

THORE

Milton Ruiz Maldonado Cédula de identidad: 1714409065

CESIÓN DE DERECHOS DE AUTOR

Yo, Jonathan Fernando Arauz Gallegos, con documento de identificación N° 1723263768, manifiesto mi voluntad y cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor/es del trabajo de grado/titulación intitulado: *"IDENTIFICACION DE FALLAS EN SISTEMAS ELECTRICOS DE POTENCIA BASADO EN EL RECONOCIMIENTO DE PATRONES"*, mismo que 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 aplicación a lo determinado en la Ley de Propiedad Intelectual, en mi condición de autor me reservo los derechos morales de la obra antes citada. En concordancia, suscribo este documento en el momento que hago entrega del trabajo final en formato impreso y digital a la Biblioteca de la Universidad Politécnica Salesiana.

Firma

Nombre: Jonathan Fernando Arauz Gallegos Cédula: 1723263768 Fecha: Quito, febrero 2020

1.INDICE GENERAL

Re	esumen
1.	Introducción2
2.	Machine Learning
	2.1 Aprendizaje no supervisado (Unsupervised Learning) 4
	2.2 Aprendizaje Supervisado (Supervised Learning) 4
	2.3 K-Vecinos más cercanos (K-Nearest Neighbors) (KNN) 4
3.	Planteamiento y Solución del Problema
4.	Método Propuesto
	4.1 Generación de la señal de entrada
	4.2 Normalización de los datos de entrada
	4.3 Vecinos más cercanos (KNN)7
5.	Resultados
	5.1 Fallas línea tierra (LG)
	5.2 Fallas línea tierra (LLG) 10
	5.3 Fallas línea línea tierra (LLLG)
6.	Conclusiones 15
7.	Referencias
8.	Estado del Arte 17

2.INDICE DE FIGURAS

Figura 1. Red de Transmisión Eléctrica.	3
Figura 2. Ejemplo IEEE 14 Barras usado en el estudio	6
Figura 3. Señal de entrada para análisis	6
Figura 4. Señal de entrada normalizada	7
Figura 5. Señal de falla tipo LG con R=0	8
Figura 6. Resultado de la identificación de la falla en la fase C	9
Figura 7. Señal de falla tipo LG con R=10	9
Figura 8. Resultado de la identificación de la falla en la fase C	9
Figura 9. Señal de falla tipo LG con R=100	10
Figura 10. Resultado de la identificación de la falla en la fase C	10
Figura 11. Señal de falla LLG con R=0.	10
Figura 12. Resultado de la identificación de la falla en la fase A y fase C.	. 11
Figura 13. Señal de falla LLG con R=10.	. 11
Figura 14 Resultado de la identificación de la falla en la fase A y fase C.	. 11
Figura 15. Señal de falla LLG con R=100.	12
Figura 16. Resultado de la identificación de la falla en la fase A y fase C.	12
Figura 17. Señal de falla LLLG con R=0	13
Figura 18 Resultado de la identificación de la falla en la fase A, fase B y fase C.	13
Figura 19. Señal de falla LLLG con R=10	13
Figura 20. Resultado de la identificación de la falla en la fase A, fase B y fase C	13
Figura 21.Señal de falla LLLG con R=100	. 14
Figura 22. Resultado de la identificación de la falla en la fase A, fase B y fase C	. 14
Figura 23. Indicadores del Estado del Arte	. 19
Figura 24. Resumen del Estado del Arte	. 19
Figura 25. Indicadores del Estado del Arte.	20

3.INDICE DE TABLAS

Tabla 1. Algoritmo para la identificación de fallas en sistemas eléctricos de potencia basado en el	
reconocimiento de patrones	. 8
Tabla 2. Tabla de distancias entre vectores de falla LG con R=0	. 9
Tabla 3. Tabla de distancias entre vectores de falla LG con R=10	. 9
Tabla 4. Tabla de distancias entre vectores de falla LG con R=100 1	10
Tabla 5. Tabla de distancias entre vectores de falla LLG con R=0 1	11
Tabla 6. Tabla de distancias entre vectores de falla LLG con R=10 1	11
Tabla 7. Tabla de distancias entre vectores de falla LLG con R=100 1	12
Tabla 8.Tabla de distancias entre vectores de falla LLLG con R=0 1	13
Tabla 9. Tabla de distancias entre vectores de falla LLLG con R=10 1	13
Tabla 10. Tabla de distancias entre vectores de falla LLLG con R=100 1	14

Tabla 11.Tabla de distancias entre vectores de falla LLLG con R=100	14
Tabla 12. Precisión del modelo para cada tipo de falla.	14
Tabla 13.Matriz del Estado del Arte	17

IDENTIFICACIÓN DE FALLAS EN SISTEMAS ELÉCTRICOS DE POTENCIA BASADO EN EL RECONOCIMIENTO DE PATRONES

Resumen

En el presente artículo se realiza la identificación de fallas eléctricas en las líneas de transmisión de un Sistema de potencia mediante un algoritmo de Machine Learning, usando el algoritmo de los vecinos más cercanos que será implementado en Matlab y PSCAD. El machine learning es cada vez más aplicado a los sistemas eléctricos debido a que ayudan al estudio de señales eléctricas y desarrollo de diversas aplicaciones las cuales se van haciendo indispensables en la actualidad. De esta forma se busca que los sistemas de potencia con equipos de comunicación avanzados entreguen una mejor calidad de energía garantizando que en una desconexión por falla esta sea resuelta en el menor tiempo posible y el identificar el tipo de falla haga que se tomen los correctivos necesarios para mantener al sistema estable, el conocer qué tipo de falla es la que se produce no es solo importante a nivel de un sistema de alta potencia, también lo es a nivel industrial para siempre garantizar la calidad de la energía.

Los patrones de falla son adquiridos desde un sistema inicial del cual se toman patrones de identificación de corriente para posteriormente realizar la simulación de casos de fallas en líneas de transmisión.

Palabras clave: identificación de fallas, Machine Learning, patrones de identificación, algoritmos vecinos más cercanos

Abstract

This article identifies electrical failures in the transmission lines of a Power System using a Machine Learning algorithm, using the algorithm of the neighbors that closest will be implemented in Matlab and PSCAD. Machine learning is increasingly applied to electrical systems because they help the study of electrical signals and the development of various applications which are going to become indispensable today. In this way, the power systems with advanced communication equipment are sought to deliver a better quality of energy ensuring that in a disconnection due to failure this sea is resolved in the shortest possible time and the identifier the type of fault causes the necessary corrective measures to be taken to keeping the system stable, knowing what type of fault is the one that occurs is not only important at a high power system level, it is also an industrial level to always control the quality of energy.

Failure patterns are acquired from an initial system from which current identification patterns are taken to subsequently simulate cases of transmission line failures.

Keywords: fault identification, Machine Learning, identification patterns, nearest neighboring algorithms

1. Introducción

En la actualidad los Sistemas Eléctricos de Potencia se han desarrollado a tal nivel que han evolucionado a redes eléctricas inteligentes que permite mantener una buena calidad de energía, en sistemas eléctricos de transmisión al ser de gran extensión el rendimiento que se tiene al ser perturbados por una falla afecta directamente la fiabilidad del Sistema Eléctrico de Potencia (SEP)[1], por lo tanto el uso de y aprovechamiento de las tecnologías de comunicaciones en estos sistemas es indispensable porque permite la comunicación entre los abonados del sistema eléctrico y las generación. centrales de esta comunicación debe ser veloz permitiendo que los centros de cómputo procesen la información recibida, se realice un monitoreo y dando una respuesta al problema.(1) La transmisión de esos datos para el monitoreo se realizan de varias formas como fibra óptica, transmisión satelital como formas de comunicación los que buscan que sean de forma óptima hasta en los eventos más críticos que puedan enfrentar, es por eso que para disminuir la cantidad de información necesaria para identificar una falla en el sistema de potencia se utilizaran patrones de fallas en líneas de transmisión ya que estos se almacenan en una base de datos v no reducida es necesaria la información de monitoreo desde dos puntos, haciendo que sea necesario solamente la información desde un punto de control[2][3].

La rápida identificación de las fallas es uno de los aspectos más relevantes ya que aportan a que se realicen trabajos de reparación más rápidos y en la toma de decisiones para las medidas a tomar , identificar los puntos débiles en las líneas de transmisión [4][5].

Al momento que se producen estas fallas las formas de onda comunes se

alteran y muestran una alteración de la forma de onda con respecto a la forma de onda común, estas formas de onda alteradas pueden ser tomadas y usadas para la detección e identificación de las fallas producidas [6]. Las formas de onda alteradas también pueden ser usadas para tomar patrones los cuales serán usados para el estudio apoyándose técnica de Aprendizaje en la Automático o Machine Learning (ML). considerado ML como es una herramienta que usa cálculos suaves para la identificación de fallas, de este modo es posible entrenar a la maquina con patrones reales existentes tomados de una parte del SEP o también de simuladores que generan estos patrones basados en la realidad.

Obtener las características apropiadas para implementar un algoritmo de ML eficiente y que permita realizar la identificación en todos los casos de estudio propuestos, una gran ventaja de este método es la baja sensibilidad a los cambios en los parámetros efectivos como son la magnitud y la dirección de la falla, el ángulo de inicio de la falla y la resistencia de falla por mencionar los más relevantes. Además, los datos tomados son de un solo extremo de la línea lo que hace que ya no sea necesario tomar datos de ambos extremos lo que hace que ya no sea necesario la transmisión de una señal y sincronizar los datos de ambos extremos[7].

Las herramientas mencionadas anteriormente tienen limitación en la dimensión de entrada de los datos que ocasiona un gran problema cuando se utiliza una gran cantidad de datos, prácticamente no son adecuados para realizar un ML con patrones.

Por otro lado, la reducción de los patrones con el uso de transformaciones lineales puede resultar con la





eliminación de información que sea útil, es por eso que los patrones resultantes de una Transformada de Wavelet [19] o una Transformada de Fourier [10] pueden ocultar algunos datos útiles.

Se propone un enfoque a utilizar el ML para identificar las fallas en las líneas de transmisión, los patrones usados son extraídos de los datos de corrientes de un solo extremo de las líneas de transmisión del SEP que se usa como ejemplo, el modelo IEEE 14 Bus-Bar, los datos obtenidos no requieren de un rango de muestreo muy alto, además la característica depende en gran medida del lugar donde se encuentra de 1a falla. como mencionó se anteriormente las principales características no tienen gran impacto en los patrones.

En este estudio se toma una gran cantidad de datos ya que se toma una muestra de cada tipo de falla que puede suceder como las L-G, L-L o L-L-G. Usando una gran cantidad de datos para identificar la falla se implementa la herramienta robusta de k –Vecino más Cercano (KNN), el cual es un algoritmo de aprendizaje supervisado, que permite el uso de patrones para realizar una identificación y clasificación del tipo de falla producido en una línea de transmisión del SEP [6], [20].

Para la generación de patrones de entrenamiento y falla, los cuales consisten en diferentes combinaciones de escenarios de fallas en el sistema de estudio. Las simulaciones se realizan en el software PSCAD/EMTDC. Seguido de la simulación se usa el método propuesto utilizando los patrones de entrenamiento y prueba obtenidos y presentar los resultados de las pruebas para la identificación de las fallas.

2. Machine Learning

ML es una parte científica del ámbito de la inteligencia artificial aprender en este plano significa el poder identificar, clasificar o aprender características diversas en un conjunto complejo de datos, tal cantidad de datos es imposible de analizar para una persona por lo cual se utiliza las máquinas para una solución correcta del problema a tratar.

Con el gran progreso de las máquinas, se han hecho necesarias el desarrollo y

aprendizaje de técnicas las cuales aprovechen la potencia y mayor capacidad de trabajo, las cuales son implementados en muchos campos del desarrollo de ingeniería.

Existen varias clasificaciones de ML debido a las tecnologías y sus nuevos algoritmos que son los siguientes:

2.1 Aprendizaje no supervisado (Unsupervised Learning)

En los problemas de aprendizaje no supervisado, no se contiene un conjunto de datos, este a su vez comienza la recolección de datos y el algoritmo se encarga de manera independiente una clasificación y comienza una clasificación de acuerdo a sus características.

Algunos de los algoritmos comprendidos en este tipo de aprendizaje son:

- Agrupamiento de K-Medias (K-Means Clustering)
- Redes Neuronales (Neuronal Network)
- Agrupamiento Jerárquico (Hierarchical Clustering)

Por mencionar varios, ya que estos tienen sus clasificaciones que tienen enfoques más exactos pero la misma base del autoaprendizaje.

2.2 Aprendizaje Supervisado (Supervised Learning)

En los problemas de aprendizaje supervisado, comenzamos el análisis con un conjunto de datos que contiene ejemplos de entrenamiento con etiquetas correctas asociadas. El algoritmo aprenderá la relación entre los datos y sus etiquetas y aplicará esa relación aprendida para clasificar datos completamente nuevos que la máquina no haya vista antes.

Algunos de los algoritmos comprendidos en este tipo de aprendizaje son:

- Regresión Lineal (Linear Regression)
- K Vecinos más Cercanos (K-Nearest Neighbors) (KNN)
- Máquina de Vector de Soportes (Support Vector Machine) (SVM)
- Árbol de decisión para Clasificación (Decision Tree Classification)

2.3 K-Vecinos más cercanos (K-Nearest Neighbors) (KNN)

KNN es uno de los algoritmos más simples y no por eso deja de ser una potente herramienta que puede dar resultados muy competitivos. Es uno de los más utilizados para el reconocimiento de patrones, debido a esto sería el algoritmo a utilizarse para el estudio de identificación de fallas eléctricas mediante patrones en líneas de transmisión.

Es un clasificador e identificador robusto y versátil que a menudo se usa como un punto de referencia para clasificadores más complejos como las ANN y SVM. A pesar de su simplicidad, KNN puede superar a los clasificadores más potentes.

Este algoritmo consiste en seleccionar un valor de K. Al momento del análisis los K datos más cercanos al valor que se desea predecir será la solución.

Acá lo importante es seleccionar un valor de K acorde a los datos para tener una mayor precisión en la identificación.

Las ventajas que tiene este algoritmo son las siguientes:

No hace suposiciones explícitas sobre la forma funcional de los datos, evitando los peligros de la distribución subyacente de los datos.

Algoritmo simple, para explicar, comprender e interpretar.

Alta precisión (relativa). Es bastante alta pero no competitiva en comparación con modelos de aprendizaje mejor supervisados.

Insensible a los valores atípicos. La precisión puede verse afectada por el ruido o las características irrelevantes.

Las desventajas de este algoritmo son las siguientes:

El algoritmo aprende no explícitamente un modelo, en su lugar, elige memorizar las instancias de capacitación que utilizan se posteriormente como conocimiento. Esto supone que, solo cuando se realiza una consulta a nuestra base de datos, es decir cuando le pedimos que prediga una etiqueta dada una entrada, el algoritmo usará las instancias de entrenamiento para entregar una respuesta.

Computacionalmente costoso, porque el algoritmo almacena todos los datos de entrenamiento. Pero con las tecnologías computacionales actuales va quedando de lado este inconveniente ya que cada vez son más robustas y rápidas en procesamiento.

Requisito de memoria alta. Almacena todos (o casi todos) los datos de entrenamiento.

Teniendo en cuenta todos los puntos de vista expuestos se toma la decisión de realizar el estudio de identificación de patrones mediante el algoritmo de los vecinos más cercanos (KNN).

3. Planteamiento y solución del problema

Como objetivo se desea determinar la identificación de fallas en líneas de transmisión para un sistema eléctrico de potencia, haciendo que los datos necesarios para su identificación sean los más adecuados posibles en la base de datos de patrones, que permitan un alto porcentaje en la identificación. El problema que se planea resolver es plantear un método diferente y darle un enfoque nuevo que pueda aportar con otros estudios, ya que el método propuesto es usado comúnmente para localizar fallas, lo que permitiría que las tecnologías de comunicación que se usen puedan manejar una base de datos consolidada y sin muchos recursos.

4. Método propuesto

El método propuesto permitirá tomar la señal de corriente a analizar, ya que, al ser identificación por patrones, no se limitaciones amplitud, tendría de duración de falla como los limitantes más visibles, ya que se hace un tratamiento previo a los datos obtenidos de la señal. Los datos de señales y patrones de fallas serán tratados como bases de datos para evitar redundancia entre la información de las señales, y por consecuencia tener una mejor condición para su análisis.

4.1 Generación de la señal de entrada

Para el estudio se utilizará el software PSCAD y el ejemplo IEEE 14 BUS BAR. Este consiste en un sistema de potencia de 14 barras a un voltaje de 138kV a 60Hz, del cual se realizará simulación de diferentes tipos de fallas como son Línea-tierra (LG), Línea-Línea-tierra (LLG), Línea-Líneatierra (LLLG), se obtendrá las señales de corriente necesarias generando la falla en las líneas de transmisión del sistema obteniendo 75 casos diferentes de fallas, por la combinación de realizar fallas y toma de señales desde diferentes puntos del sistema de potencia. Después de cada simulación las simulaciones serán gravadas en un archivo el cual será tomado por Matlab.



Figura 2. Ejemplo IEEE 14 Barras usado en el estudio.

4.2 Normalización de los datos de entrada

La normalización de datos es una parte fundamental para evitar efectos negativos en los resultados deseados, necesario para los datos de las señales al ser tomadas como bases de datos, por eso es necesario antes del análisis y comparación de las señales hacer la normalización para asegurar que los datos estén organizados, minimizar la cantidad de datos duplicados y evitar que valores de rango muy altos puedan solapar a los demás datos, haciendo que el resultado sea una base de datos de acceso fácil y rápido.

Se entiende que la normalización es la operación de transformar un conjunto de valores o bases de datos de una magnitud determinada y convertirlos en otro, para que todos esos conjuntos de datos se encuentren en una misma escala determinada y permitir un uso más fácil de estos datos.



Figura 3. Señal de entrada para análisis.

Para esta operación es posible normalizar estos valores en un intervalo [0,1] aplicando la siguiente transformación:

$$e_k = \frac{a_k - min}{max - min},$$
 (1)
con *i* = 1 hasta *N*

Donde: a_i es el valor a transformar

min es el valor mínimo del conjunto de datos

max es el valor máximo del conjunto de datos

 e_k es el valor normalizado del conjunto de datos

N número total de datos

Esta operación será aplicada a cada dato de las señales tratadas y a los patrones de falla obtenidos para que se encuentren en una misma escala que servirá posteriormente tal como se muestra en la figura (4).



4.3 Vecinos más cercanos (KNN)

KNN es un algoritmo simple y de alta preferencia del aprendizaje supervisado, basada en un conjunto de entrenamiento y patrones, debido a que permite seguir con una estrategia de poco tiempo de esfuerzo al entrenamiento y un esfuerzo completo al momento de la predicción que sirve para estimar la función más cercana de los predictores x por cada clase C que se tiene, haciendo la estimación de que un elemento x pertenezca a una clase con la información proporcionada.

Para el reconocimiento de patrones, KNN se usa como un tipo de "Lazy Learning" donde los datos se aproximan solo localmente. Los ejemplos de entrenamiento usados son planteados como vectores en un espacio característico el cual para cada uno de sus ejemplos se describe como términos de p atributos que son considerados con q clases para la identificación de los patrones. Los valores de los atributos de un i - esimo ejemplo donde el valor de *n* se encuentra $i \leq 1 \leq n$ y son representados por el vector p –

 $dimensional X_i = (X_{1i}, X_{21}, \dots, X_{pi}) \in X.$

Estos vectores almacenados, patrones, son comparados con el vector de señal a analizar, con lo cual se determinan los vecinos más cercanos y estos son asignados como válidos correspondiente distancia. а su generalmente usada la distancia euclidiana dada por la ecuación:

$$d(x_{i}, x_{j}) = \sqrt{\sum_{r=1}^{p} (x_{ri} - x_{rj})^{2}}$$
(2)

La fase inicial de entrenamiento está basada en el almacenamiento de los vectores característicos de las clases de eiemplos de entrenamiento, en la segunda fase la cual se ingresa el vector de ejemplo del cual no se conoce su clase es presentado como un vector del espacio característico. Se calcula la distancia ente los vectores almacenados como patrones y el nuevo vector, en este caso la señal de corriente a analizar y son seleccionados los K ejemplos más cercanos, en este estudio k = 1, ya que el interés es encontrar la distancia mínima punto a punto minimizando el duplicar datos de forma innecesaria, este nuevo ejemplo resultado es comparado con la clase que más se acerque teniendo la menor distancia.

Con el objetivo de tratar de disminuir la dependencia del costo computacional del modelo KNN respecto al número de casos de fallas almacenados, existe métodos de reducción de ficheros tales como técnicas de edición y técnicas de condensación, para hacer los datos de uso efectivo.

En el estudio se aplicará la Edición de Wilson que propone someter a prueba a cada caso inicial, para ello se compara cada clase verdadera con el resto de casos propuesto por KNN, de cierto modo el método tiene una analogía de validación leave-one-out, para que sea posible mostrar la respuesta deseada.

 Tabla 1. Algoritmo para la identificación de fallas en sistemas eléctricos de potencia basado en el reconocimiento de patrones.

Paso	Adquisición de datos del Sistema			
1	de Potencia Eléctrico			
	Búsqueda del inicio de la falla generada			
Paso 2	$J = \sum_{i=1}^{k-1} (x_i - x_1^{k-1})^2$			
	$+\sum_{i=k}(x_i-x_k^N)^2$			
Normalización de la señal de				

entrada

$$\frac{\text{Paso}}{3} \quad e_k = \frac{a_k - min}{max - min}$$

 $\operatorname{con} i = 1$ hasta N

Normalización de los patrones

Paso
$$e_k = \frac{a_k - min}{max - min}$$
,

con i = 1 hasta N

Calculo de las distancias entre vectores usando KNN

Paso 5
$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{p} (x_{ri} - x_{rj})^2}$$

Matriz de distancias entre vectores

	$DD_{i,j} = d(x_i, x_j)$
Paso	Aplicación del método de
6	validación Edición de Wilson
Paso	Apólicie y optrogo do recultados
7	Analisis y entrega de resultados

5. Resultados

El método KNN propuesto es probado en varios casos de fallas para probar el rendimiento del modelo. Algunos de los resultados se presentan de acuerdo al tipo de falla.

5.1 Fallas línea tierra (LG)

Todas las fallas son generadas en las líneas de transmisión, las fallas LG tienen una probabilidad de 80-85% de producirse, por eso es tan importante que se pueda identificar una falla para que sus respectivas protecciones actúen correctamente.

El método propuesto es probado para distintos tipos de fallas a tierra como AG, BG y CG, la resistencia de falla utilizada es $R = 0\Omega$, $R = 10\Omega$, $R = 10\Omega\Omega$ y la falla se produce a t = 0.2s con una duración de t = 0.02s.

En la figura (5) se muestra la señal generada de corriente con $R = 0\Omega$ en uno de las barras del sistema desde PSCAD.



En la figura (6) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.173491s en identificar la falla de la señal de entrada.

Como resultado tenemos que en la fase A tiene un ligero cambio de amplitud por la falla generada y no es reconocida como falla ya que no reúne las características para ser reconocida, en la fase B se tiene el mismo caso la señal es afectada peor no tiene las características para reconocer como falla y en la fase C se puede observar que la falla es



Figura 6. Resultado de la identificación de la falla en la fase C.

reconocida ya que el patrón de falla es similar a la señal de entrada tomada en el análisis.

También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (2).

Tabla 2. Tabla de distancias entre vectores de falla $LG \operatorname{con} B=0$

		Vectores de Señal de		
		entrada		
		A B C		
T 7 4	A	0.2666	0.3063	0.0047
Vectores de Falla	B	0.1451	0.1774	0.0057
ut l'alla	С	0.095	0.1200	0.0044

En la figura (7) se muestra la señal generada de corriente con $R = 10\Omega$ en uno de las barras del sistema desde PSCAD.



En la figura (8) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.145831s en identificar la falla de la señal de entrada.



Figura 8. Resultado de la identificación de la falla en la fase C.

Como resultado tenemos que en la fase A tiene un ligero cambio de amplitud por la falla generada y no es reconocida como falla ya que no reúne las características para ser reconocida, en la fase B se tiene el mismo caso la señal es afectada peor no tiene las características para reconocer como falla y en la fase C se puede observar que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada tomada en el análisis.

También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (3).

EG con R=10.				
	Vectores de Señal de entrada			
		Α	В	С
Vectores	Α	0.3406	0.2742	0.0050
de Falla	В	0.3090	0.2417	0.0062
ue ralla	С	0.2130	0.1599	0.0036

Tabla 3. Tabla de distancias entre vectores de fallaLG con R=10.

En la figura (9) se muestra la señal generada de corriente con $R = 10\Omega$ en uno de las barras del sistema desde PSCAD.



En la figura (10) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.0278141s en identificar la falla de la señal de entrada.



Como resultado tenemos que en la

fase A tiene un ligero cambio de amplitud por la falla generada y no es reconocida como falla ya que no reúne las características para ser reconocida, en la fase B se tiene el mismo caso la señal es afectada peor no tiene las características para reconocer como falla y en la fase C se puede observar que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada tomada en el análisis.

También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (4).

Se realiza pruebas de fallas generadas en distintas líneas y tomando la señal en diferentes barras del sistema, con repeticiones de 5 veces de cada caso para confirmar el resultado.

LO COII K-100.					
		Vectores de Señal de entrada			
		Α	В	С	
X 7 4	A	0.2041	0.1239	0.0054	
vectores de Falla	B	0.0973	0.0380	0.0052	
ut rana	С	0.0649	0.0213	0.0038	

Tabla 4. Tabla de distancias entre vectores de falla

En la matriz generada por las distancias, la diagonal principal muestra las distancias calculadas entre los vectores. Teniendo las de mínima distancia más cercanas a cero o valores menores a 0.01. El valor de 0.01 se adquirió mediante las pruebas de todas las fallas, llegando a ese valor admisible para dar como positivo el detectar la falla. Este concepto se aplicará a todos los casos de fallas generados AG, BG y CG.

5.2 Fallas línea línea tierra (LLG)

El método propuesto también se puede realizar pruebas para la identificación de fallas de tipo LLG, se realiza en las mismas condiciones que las fallas LG, con R = 0, $R = 10\Omega$, $R = 100\Omega$ y la falla se produce a t = 0.2s con una duración de t = 0.02s.

En la figura (11) se muestra la señal generada de corriente en uno de las barras del sistema desde PSCAD.



En la figura (12) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.044709s en identificar la falla de la señal de entrada.



Figura 12. Resultado de la identificación de la falla en la fase A y fase C.

Como resultado tenemos que en la fase A se observa que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada, en la fase B se tiene la señal es afectada pero no tiene las características para reconocer como falla y en la fase C se puede ver que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada tomada en el análisis.

También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (5).

 LLG con R=0.

		Vectores de Señal de entrada		
		A B C		
Vectores	Α	0.0041	0.4367	0.0870
do Follo	В	0.0053	0.2870	0.0183
ue ralla	С	0.0061	0.2214	0.0089

En la figura (13) se muestra la señal generada de corriente en uno de las barras del sistema desde PSCAD.

En la figura (14) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.007952s en identificar la falla de la señal de entrada.



Como resultado tenemos que en la fase A se observa que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada, en la fase B se tiene la señal es afectada pero no tiene las características para reconocer como falla y en la fase C se puede ver que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada tomada en el análisis.



en la fase A y fase C.

También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (6).

Tabla 6. Tabla de distancias entre vectores de falla
LLG con R=10.

		Vectores de Señal de entrada		
		A B C		
Mastaria	Α	0.0043	0.4196	0.0192
do Falla	В	0.0064	0.2700	0.0052
ue ralla	С	0.0068	0.2069	0.0033

En la figura (15) se muestra la señal generada de corriente en uno de las barras del sistema desde PSCAD.



En la figura (16) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.546382s en identificar la falla de la señal de entrada.



falla en la fase A y fase C.

Como resultado tenemos que en la fase A se observa que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada, en la fase B se tiene la señal es afectada pero no tiene las características para reconocer como falla y en la fase C se puede ver que la falla es reconocida ya que el patrón de falla es similar a la señal de entrada tomada en el análisis.

También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (7).

Tabla 7. Tabla de distancias entre vectores de falla
LLG con R=100.

		Vecto	Vectores de Señal de entrada										
		Α	В	С									
Mastaria	Α	0.0035	0.2059	0.0075									
do Falla	В	0.0056	0.0973	0.0052									
ue ralla	С	0.0056	0.0593	0.0034									

Se realiza pruebas de fallas generadas en distintas líneas y tomando la señal en diferentes barras del sistema, con repeticiones de 5 veces de cada caso para confirmar el resultado. Como resultado nos indica que la falla generada es en la fase B y fase C.

En la matriz generada por las distancias, la diagonal principal muestra las distancias calculadas entre los vectores. Teniendo las de mínima distancia más cercanas a cero o valores menores a 0.01. El valor de 0.01 se adquirió mediante las pruebas de todas las fallas, llegando a ese valor admisible para dar como positivo el detectar la falla. Este concepto se aplicará a todos los casos de fallas generados ABG, BCG y ACG

5.3 Fallas línea línea línea tierra (LLLG)

El método propuesto también se puede realizar pruebas para la identificación de fallas de tipo LLLG, se realiza en las mismas condiciones que las fallas LG, $R = 0\Omega$, $R = 10\Omega$, $R = 100\Omega$ y la falla se produce a t = 0.2s con una duración de t = 0.02s.

En la figura (17) se muestra la señal generada de corriente en uno de las barras del sistema desde PSCAD.



En la figura (18) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.087056s en identificar la falla de la señal de entrada.

Como resultado se tiene que tanto en la fase A, fase B y fase C tiene características similares a los patrones de falla y por lo tanto dando como respuesta positiva la identificación.



También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (8).

 Tabla 8.Tabla de distancias entre vectores de falla

 LLLG con R=0.

		Vectores de Señal de entrada										
		Α	В	С								
	Α	0.0047	0.0162	0.0194								
vectores	В	0.0062	0.0035	0.0054								
ue ralla	С	0.0076	0.0041	0.0035								

En la figura (19) se muestra la señal generada de corriente en uno de las barras del sistema desde PSCAD.



En la figura (20) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.657010s en identificar la falla de la señal de entrada.



Como resultado se tiene que tanto en la fase A, fase B y fase C tiene características similares a los patrones de falla y por lo tanto dando como respuesta positiva la identificación. También se puede indicar el resultado mediante las distancias generadas por el modelo como en la tabla (9).

Tabla 9.	Tabla de distancias entre vectores de fa	alla
	LLLG con R=10.	

		Vecto	Vectores de Señal de entrada											
		Α	В	С										
Mastanaa	Α	0.0054	0.0053	0.0068										
Vectores de Falla	В	0.0054	0.0049	0.0059										
	С	0.0058	0.0057	0.0042										

En la figura (21) se muestra la señal generada de corriente en uno de las barras del sistema desde PSCAD.



En la figura (22) se muestra la salida del modelo realizado y su respuesta confirmando en qué fase sucedió la falla, el modelo tuvo 3.524201s en identificar la falla de la señal de entrada.



Figura 22. Resultado de la identificación de la falla en la fase A, fase B y fase C.

Como resultado se tiene que tanto en la fase A, fase B y fase C tiene características similares a los patrones de falla y por lo tanto dando como respuesta positiva la identificación. También se puede indicar el resultado

mediante las distancias generadas por el modelo como en la tabla (10).

Tabla 10. Tabla de distancias entre vectores de
falla LLLG con R=100.

		Vectores de Señal de entrada										
		Α	В	С								
Mastanaa	Α	0.0052	0.0053	0.0068								
Vectores de Falla	В	0.0053	0.0037	0.0052								
	С	0.0055	0.0050	0.0036								

De estos resultados también se puede obtener una matriz de confusión la que nos indica los falsos positivos y falsos negativos generados en los casos de prueba y así determinar la precisión del modelo propuesto como se muestra en la tabla (11).

		$1 \text{ con } \mathbf{K} = 1$	00.	
		Val	or Pred	icho
		Falla	Falla	
		LG	LLG	LLG
	Falla LG	124	9	1
Valor Real	Falla LLG	7	140	3
	Falla LLG	0	4	46

Tabla 11.Tabla de distancias entre vectores de fallaLLLG con R=100.

Se tiene que para una falla tipo LG se tiene 135 pruebas comprendidas entre fallas AG, BG y CG. Para una falla tipo LLG se tiene 150 pruebas comprendidas entre fallas ABG, ACG y BCG. Para una falla tipo LLLG se tiene 50 pruebas de la falla ABCG.

Con esos resultados se puede determinar que el modelo utilizado puede ser de alta precisión entregando buenos resultados en la mayoría de los casos tal como se muestra en la tabla (12).

 Tabla 12. Precisión del modelo para cada tipo de falla

Tipo de Falla	Número de Casos	Precisión (%)
LG	134	92,54
LLG	150	93,33
LLLG	50	92

Se realiza los mismos casos para una diferente impedancia de falla, lo que causa que la amplitud de las señales disminuyan, pero obteniendo resultados similares de precisión al tener una impedancia de falla cero.

6. Conclusiones

En la presente investigación se pudo determinar que a pesar de que es un algoritmo simple puede entregar un gran resultado y al usar el algoritmo de KNN es muy flexible y se lo puede utilizar de varias formas, pero en un caso particular usado para identificar con patrones guardados en bases de datos, esto hace que se deje abierto a futuros estudios a que se realice con diferentes bases de datos de patrones en casos deseados.

Se pudo demostrar que es posible realizar un reconocimiento a través de patrones, a pesar de que no es una forma común, pero dando un buen resultado en el que incluso se pudo mostrar que el modelo puede demorar alrededor de 3s en identificar cada tipo de falla, además se pudo determinar que la amplitud de las señales no sería un problema después de normalizar la señal.

También se pudo determinar al variar la impedancia de falla que disminuyen los casos de prueba, pero también permitió ver un radio de afectación de la falla producida en las líneas de trasmisión hacia los elementos y equipos del sistema de potencia, es decir ver hasta donde llegaría la afectación del sistema alrededor de la falla producida.

7. Referencias

- M. Ruiz and E. Inga, "Despliegue óptimo de redes ópticas para comunicaciones en redes eléctricas inteligentes," *I+D Tecnológico*, vol. 15, no. 2, pp. 79–85, 2019.
- [2] M. Ruiz, P. Masache, and J. Dominguez, "High availability network for critical communications on smart grids," *CEUR Workshop Proc.*, vol.

2178, no. Ssn, pp. 13–17, 2018.

- [3] M. G. Ruiz Maldonado and E. Inga, "Asignación óptima de recursos de comunicaciones para sistemas de gestión de energía," *Enfoque UTE*, vol. 10, no. 1, pp. 141–152, 2019.
- [4] A. Zapata-Tapasco, S. Pérez-Londoño, and J. Mora-Flórez, "Método basado en clasificadores k-NN parametrizados con algoritmos genéticos y laestimación de la reactancia para localización de fallas en sistemas de distribución," *Rev. Fac. Ing.*, no. 70, pp. 220–232, 2014.
- [5] M. M. Saha, J. Izykowski, and E. Rosolowski, "Fault Location on Power Networks," *Power Syst.*, vol. 48, pp. 1–422, 2010.
- [6] A. Yadav and A. Swetapadma, "Fault analysis in three phase transmission lines using k-nearest neighbor algorithm," 2014 Int. Conf. Adv. Electron. Comput. Commun. ICAECC 2014, pp. 1–5, 2015.
- [7] R. Mardiana, H. Al Motairy, and C. Q. Su, "Ground fault location on a transmission line using highfrequency transient voltages," *IEEE Trans. Power Deliv.*, vol. 26, no. 2, pp. 1298–1299, 2011.
- [8] Y. Sheng and S. M. Rovnyak, "Decision tree-based methodology for high impedance fault detection," *IEEE Trans. Power Deliv.*, vol. 19, no. 2, pp. 533–536, 2004.
- [9] S. Osowski and R. Salat, "Fault location in transmission line using hybrid neural network," *COMPEL Int. J. Comput. Math. Electr. Electron. Eng.*, vol. 21, pp. 18–30, 2002.
- [10] T. Bouthiba, "Fault location in EHV transmission lines using artificial neural networks," *Int. J. Appl. Math. Comput. Sci*, vol. 14, pp. 69–78, 2004.

- J. Ezquerra, V. Valverde, A. [11] Mazón, I. Zamora, and J. J. Zamora, "Field programmable gate array implementation of a location fault system in transmission lines based on artificial neural networks," IET Gener. Transm. Distrib., vol. 5, no. 2, pp. 191–198, 2011.
- [12] R. Salat and S. Osowski, "Accurate fault location in the power transmission line using support vector machine approach," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 979–986, 2004.
- [13] V. Malathi, N. S. Marimuthu, and S. Baskar, "Intelligent approaches using support vector machine and extreme learning machine for transmission line protection," *Neurocomputing*, vol. 73, no. 10, pp. 2160–2167, 2010.
- [14] U. B. Parikh, B. Das, and R. Maheshwari, "Fault classification technique for series compensated transmission line using support vector machine," *Int. J. Electr. Power Energy Syst.*, vol. 32, no. 6, pp. 629–636, 2010.
- [15] Y. Zhang, J. Chen, Q. Fang, and Z. Ye, "Fault analysis and prediction of transmission line based on fuzzy K-Nearest Neighbor algorithm," 2016 12th Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov. ICNC-FSKD 2016, pp. 894–899, 2016.
- [16] M. J. Reddy and D. K. Mohanta, "A Wavelet-neuro-fuzzy Combined Approach for Digital Relaying of Transmission Line Faults," *Electr. Power Components Syst.*, vol. 35, no. 12, pp. 1385–1407, 2007.
- [17] M. J. B. Reddy and D. K. Mohanta, "Performance evaluation of an adaptivenetwork-based fuzzy inference system approach for location of

faults on transmission lines using Monte Carlo simulation," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 4, pp. 909–919, 2008.

- [18] J. Sadeh and H. Afradi, "A new and accurate fault location algorithm for combined transmission lines using Adaptive Network-Based Fuzzy Inference System," *Electr. Power Syst. Res.*, vol. 79, no. 11, pp. 1538–1545, 2009.
- [19] S. Ekici, S. Yildirim, and M. Poyraz, "A transmission line fault locator based on Elman recurrent networks," *Appl. Soft Comput.*, vol. 9, no. 1, pp. 341–347, 2009.
- [20] G. Morales-España, J. Mora-Flórez, and G. Carrillo-Caicedo, fault location "A complete formulation for distribution systems using the k-nearest neighbors for regression and classification," 2010 IEEE/PES Transm. Distrib. Conf. Expo. Lat. Am. T D-LA 2010, pp. 810-815, 2011.

8. Estado del arte

Tabla 13. Matriz del Estado del Arte

IDENTIFICACION DE FALLAS EN SISTEMAS ELÉCTRICOS DE POTENCIA BASADO EN EL RECONOCIMIENTO DE PATRONES

		DATOS			Tŀ	EMÁTI	CA		FORMULACIÓN DEL PROBLEMA FUNCIONES OBJETIVO						RESTRICCIONES DEL PROBLEMA						PUES VER E	TAS P L PRO	PARA DBLEN	1A	SOLUCIÓN PROPUESTA					
ITEM	AÑO	TÍTULO DEL ARTÍCULO	N° DE CITAS PUBLISH OR PERISH	ALGORITMOS DE CLASIFICACION	LOCALIZACIÓN DE FALLAS	EFICIENCIA DE LOS DATOS	REDES ELÉCTRICAS DE TRANSMISIÓN	IDENTIFICACION DE FALLAS	IDENTIFICACIÓN DE FALLAS	LOCALIZACIÓN DE FALLAS	ANÁLISIS DE SEÑALES	REDES NEURONALES	SMART GRIDS	BASE DE DATOS	IMPEDANCIA DE FALLA	PRECISIÓN DE LOS PATRONES	TIEMPO DE LOCALIZACIÓN DE FALLA	COBERTURA	K VECINOS CERCANOS	VECTORES SOPORTE DE MAQUINA	LOGICA DIFUSA	ÁRBOL DE DECISIONES	REDES NEURONALES	WAVELETS	K VECINOS CERCANOS	IDENTIFICACION DE FALLAS	WAVELETS	REDES ADAPTATIVAS	REGRESION Y CLASIFICACION	
1	2019	Despliegue óptimo de redes ópticas para comunicaciones en redes eléctricas inteligentes	0			¥	æ				æ		₩					æ												
2	2018	High availability network for critical communications on smart grids	11			¥	æ				æ		æ					¥												
3	2019	Asignación óptima de recursos de comunicaciones para sistemas de gestión de energía	4			¥	¥				æ		æ					¥												
4	2014	Método basado en clasificadores k- NN parametrizados con algoritmos genéticos y la estimación de la reactancia para localización de fallas en sistemas de distribución	3	¥	¥		¥	₩	¥	æ	¥			₩	¥		¥		æ				¥		挼	¥			₩	
5	2010	Fault Location on Power Networks	599	æ	×		×	×	×						×		æ													
6	2015	Fault analysis in three phase transmission lines using k-nearest neighbor algorithm	11	æ	₩		₩	æ		₩	₩				₩			₩	*					₩	₩		₩		₩	
7	2011	Ground fault location on a transmission line using high- frequency transient voltages	39	₩	₩		₩	æ	₩	₩	₩				æ		₩	æ	₩				æ		¥				æ	

8	2004	Decision tree-based methodology for high impedance fault detection	206	₩	₩		₩	₩			æ				₩		æ	₩				¥				₩			æ
9	2002	Fault location in transmission line using hybrid neural network	13	æ	æ		æ	æ				æ				æ	æ	æ					æ						æ
10	2004	Fault location in EHV transmission lines using artificial neural networks	100	₩	₩		₩	₩		₩		₩					₩						¥					æ	₩
11	2011	Field programmable gate array implementation of a fault location system in transmission lines based on artificial neural networks	17	₩	¥		æ	₩		₩	æ	₩					₩						¥					¥	æ
12	2004	Accurate fault location in the power transmission line using support vector machine approach	173	₩	₩		₩	₩		₩	₩			₩			₩			Ħ						¥			æ
13	2010	Intelligent approaches using support vector machine and extreme learning machine for transmission line protection	72	æ	¥		æ	æ						₩			₩			¥						Æ			æ
14	2010	Fault classification technique for series compensated transmission line using support vector machine	107	₩	₩		₩	₩						₩			₩			瘷						¥			₩
15	2016	Fault analysis and prediction of transmission line based on fuzzy K- Nearest Neighbor algorithm	4	æ	₩	æ	₩	₩		₩									₩		₩					¥			₩
16	2007	A Wavelet-neuro-fuzzy Combined Approach for Digital Relaying of Transmission Line Faults	62	æ	₩	æ	₩	₩				¥									₩		₩					æ	æ
17	2008	Performance evaluation of an adaptive-network-based fuzzy inference system approach for location of faults on transmission lines using Monte Carlo simulation	93	æ	¥		æ	æ				¥									¥		¥					¥	æ
18	2009	A new and accurate fault location algorithm for combined transmission lines using Adaptive Network-Based Fuzzy Inference System	102	¥	¥	₩	₩	₩				₩		₩							¥		¥	¥				¥	æ
19	2009	A transmission line fault locator based on Elman recurrent networks	69	Ð	Ŧ	¥	₩	₩		₩		¥		æ							₩		₩						æ
20	2011	A complete fault location formulation for distribution systems using the k-nearest neighbors for regression and classification	9	₩	*		*	₩		¥					₩		¥		₩										×
			CANTIDAD:	17	17	7	20	17	3	10	9	7	3	6	6	1		7	5	3	5	1		2	3	6	1	5	16



Figura 24. Resumen del Estado del Arte



Figura 23. Indicadores del Estado del Arte



Figura 25. Indicadores del Estado del Arte.