

Partitioning of an Electrical Distribution Systems Using K-Means and DBSCAN Clustering Algorithms

Partición de una Red Eléctrica de Distribución Aplicando Algoritmos de Agrupamiento K-means y DBSCAN

 P.A. Soto¹

 J.R. Castro¹

 R.M. Reategui¹

 T.D. Castillo¹

¹Universidad Técnica Particular de Loja

E-mail: pasoto@utpl.edu.ec; jrcaastro@utpl.edu.ec; mreategui@utpl.edu.ec; tdcastillo@utpl.edu.ec

Abstract

This paper proposes the methodology to perform the partitioning of a distribution network using data clustering algorithms such as K-means and DBSCAN. The data is obtained by generating variations in the network parameters and simulating the voltage profile using OpenDSS software. The proposed methodology is implemented on standard IEEE test distribution networks of 34 and 123 node test feeder. The results show that the nodes are grouped, achieving an adequate partition of the electrical distribution network.

Index terms— Data Mining, Distribution Networks, Partition Electric, Clustering.

Resumen

En este artículo se propone la metodología para realizar la partición eléctrica de una red de distribución utilizando algoritmos de agrupamiento de datos como K-means y DBSCAN. Los datos se obtienen generando variaciones en los parámetros de la red y simulando el perfil de voltaje con el software OpenDSS. La metodología propuesta se implementa en redes de distribución estándar de prueba IEEE de 34 y 123 nodos. Los resultados muestran que los nodos se agrupan, logrando una adecuada partición de la red eléctrica de distribución.

Palabras clave— Minería de Datos, Redes de Distribución, Partición Eléctrica, Agrupación.

Recibido: 13-04-2023, Aprobado tras revisión: 14-06-2023

Forma sugerida de citación: Soto, P.; Castro, J.; Reategui, R.; Castillo, T. (2023). "Partición de una Red Eléctrica de Distribución Aplicando Algoritmos de Agrupamiento K-Means y DBSCAN". Revista Técnica "energía". No. 20, Issue I, Pp. 73-81

ISSN On-line: 2602-8492 - ISSN Impreso: 1390-5074

Doi: <https://doi.org/10.37116/revistaenergia.v20.n1.2023.572>

© 2023 Operador Nacional de Electricidad, CENACE



1. INTRODUCCIÓN

El requerimiento de consumo de energía eléctrica en todos los países del mundo aumenta cada año. La generación, transmisión y distribución de energía presenta uno de los índices más altos de crecimiento a nivel global. Debido a la rápida industrialización y el crecimiento de los sectores residencial y comercial, las de las redes de distribución eléctrica se enfrentan a problemas de saturación y envejecimiento [1]. El crecimiento y la diversificación de las redes eléctricas en los últimos años han llevado a que operen cerca de sus límites de estabilidad y seguridad, lo que incrementa significativamente la probabilidad de fallas [2].

Los sistemas de energía actuales están en una fase de transición hacia redes más extensas y complejas. El método tradicional de control centralizado enfrenta el desafío de una gran carga computacional, especialmente en el caso de redes a gran escala, debido a la enorme cantidad de información que genera [3]. Estudios señalan que la mejor estrategia para superar estos desafíos es la subdivisión de las redes eléctricas en múltiples áreas interconectadas [4].

Varias investigaciones presentan diferentes métodos para la partición de redes eléctricas de distribución [4][5][6][7][8]. Mao [5] propone tres pasos para la partición de redes eléctricas. El primero paso realiza el cálculo de la distancia eléctrica a través del método de la impedancia. En segundo paso, aplica el método de detección de comunidades para dividir la red eléctrica en zonas internas, Y el tercer paso, une las zonas utilizando un algoritmo de búsqueda.

La partición de redes eléctricas puede mejorar el desempeño del sistema de distribución. Aumenta la seguridad ante fallas o ataques externos, mejorara la confiabilidad y la calidad del servicio al reducir las interrupciones y las fluctuaciones, optimiza la eficiencia al disminuir las pérdidas y el consumo, facilita la integración de fuentes de generación distribuida y de redes inteligentes, y favorece la participación y la satisfacción de los usuarios [9]. Sin embargo, esta técnica también implica algunos desafíos, como: definir los criterios y los métodos más adecuados para cada caso, considerar la incertidumbre y la variabilidad de las condiciones del sistema, evaluar el impacto ambiental y social de la segmentación, y garantizar la coordinación y la comunicación entre las zonas de la red.

Chai [4] propone como estrategia para controlar las variaciones de voltaje, la partición de redes eléctricas de distribución basada en el método de dirección alterna de multiplicadores. La red IEEE 123 nodos es fraccionada en 4 zonas.

Métodos de optimización lineal como los presentados en [8] y métodos de análisis de sensibilidad [10] carecen en incluir todas las variabilidades de las redes eléctricas; para incluir variaciones estocásticas en la red, desbalance de cargas y múltiples sistemas de control y protecciones

el problema de segmentación necesita ser resuelto por técnicas no lineales como los presentados en [9], que usa un algoritmo de recocido simulado para resolver el problema de minimización buscando un balance entre número, tamaño y autosuficiencia de los segmentos detectados

En la operación en tiempo real de una red eléctrica, los distribuidores de este servicio a menudo confían en una serie de reglas de operación seguras y estables para mantener la operatividad del sistema eléctrico. La potencia de transmisión límite o la capacidad de transmisión total es uno de los indicadores de operación clave o más importantes dentro de una red eléctrica interconectada. En los últimos años, el desarrollo y la amplia aplicación de tecnologías de big data e inteligencia artificial han proporcionado nuevos medios técnicos para el modelado, operación y distribución de energía en redes eléctricas [11].

La minería de datos utilizando técnicas de identificación de patrones y algoritmos de aprendizaje, ha revolucionado la forma de comprender y manejar sistemas, brindando soporte en el control de procesos, operaciones y servicios a los operadores de las diferentes áreas que manejan grandes volúmenes de datos. Una técnica de minería de datos es el análisis de agrupamientos, que permite identificar grupos con características similares, a través de algoritmos matemáticos y estadísticos. En [2] se usa minería de datos para identificar la barra piloto en redes eléctricas de distribución de IEEE de 13 y 34 nodos utilizando K-Means y DBSCAN.

En el presente trabajo se utilizará los algoritmos de agrupación K-Means y DBSCAN, ampliamente utilizados en minería de datos, con el fin de llevar a cabo la partición de las redes eléctricas de distribución IEEE de 34 y 123 nodos en el entorno de desarrollo integrado R-Studio.

La estructura del presente trabajo se compone de la siguiente manera: en el Capítulo 2 se describe las redes de distribución eléctrica, en el Capítulo 3 se detalla los pasos para la implementación de los algoritmos de minería de datos, el Capítulo 4 presenta el caso de estudio junto con los resultados obtenidos utilizando con la metodología propuesta, y finalmente en el capítulo 5 se exponen las conclusiones alcanzadas.

2. REDES DE DISTRIBUCIÓN

Las empresas eléctricas, en cumplimiento de las normativas vigentes, opera sistemas eléctricos que abarcan la generación, transmisión, distribución y utilización de la energía eléctrica. Su función primordial es asegurar el suministro continuo y de calidad, llevando esta energía desde su origen hasta los usuarios finales, como un servicio público esencial.

La estabilización del voltaje a lo largo de la red de distribución es un desafío clave en la operación de

sistemas eléctricos. Estos sistemas suministran potencia activa y reactiva a los usuarios finales, cuya demanda energética varía constantemente. Esto ocasiona fluctuaciones de voltaje en todos los puntos de la red. Para controlar estas variaciones, las compañías eléctricas emplean bancos de capacitores para inyectar potencia reactiva cuando es necesario, y, de manera inversa, utilizan bancos de inductores para consumir potencia reactiva según se requiera.

El sistema de prueba denominado IEEE de 34 nodos es una red desbalanceada, contiene 34 nodos, 32 líneas o enlaces que unen los nodos, de los cuales 24 son trifásicos y 8 monofásicos. La red está conformada por líneas monofásicas, bifásicas y trifásicas, 2 reguladores, 3 capacitores, una tensión nominal de 24.9kV. El alimentador de prueba IEEE de 123 nodos posee una tensión nominal de 4.16kV, está conformado por 123 nodos de los cuales 85 son de carga y uno de alimentación, consta de 118 enlaces aéreos y subterráneos o líneas entre los nodos, 4 reguladores de voltaje para solucionar los problemas de caídas de voltaje, 4 capacitores, cargas desbalanceadas y 12 interruptores [12].

Los datos empleados se obtuvieron mediante el software de simulación de redes eléctricas OpenDSS [13], utilizando el método estadístico de distribución Monte Carlo. Se generaron 100 variaciones de los niveles de tensión en cada nodo de las redes de prueba, donde cada variación de carga se realizó entre el 95% y el 105% de su valor nominal.

Una vez conformadas las bases de datos, es necesario realizar una normalización de los datos, lo cual permite una comparación adecuada y evita que los valores más grandes afecten los resultados del análisis. Para determinar el número óptimo de agrupaciones para cada algoritmo, se empleará el método del codo, y se utilizará el coeficiente de silueta como métrica para evaluar la calidad de los grupos.

2.1. Importancia de la partición eléctrica

El objetivo de la partición de redes de distribución eléctrica es dividir la red en áreas operativamente independientes y restringir el intercambio de información solo entre los nodos que conectan áreas adyacentes [8]. Para lograr la formación de grupos, se define una medida de similitud que refleje la influencia entre los nodos de la red, y luego se ejecuta el algoritmo de agrupamiento correspondiente.

A medida que la estructura de la red de distribución se vuelve más compleja, su análisis y monitoreo se vuelven más desafiantes. La segmentación de la red puede simplificar su estructura, permitiendo una operación y controles más eficientes en cada región [14].

Para asegurar una operación y gestión segura y estable de las redes eléctricas regionales, es crucial realizar una planificación de red adecuada y viable. Con

el fin de monitorear en tiempo real el estado de operación de la red y tomar decisiones rápidas, los profesionales del sector energético suelen dividir la red en varias subregiones y administrar cada una de forma independiente, lo que mejora significativamente la velocidad de procesamiento y reduce la carga computacional [15]. El control del voltaje en toda la red puede lograrse mediante el control individual de cada grupo [16]. Los resultados obtenidos en este estudio son comparados con los hallazgos de otros autores en la literatura.

3. MINERÍA DE DATOS

En la actualidad, la mayoría de los sistemas generan enormes cantidades de información que requieren ser procesadas y analizadas de manera ordenada, utilizando herramientas informáticas para automatizar este proceso. La minería de datos se destaca como un conjunto de técnicas diseñadas para extraer información valiosa a partir de grandes volúmenes de datos, analizando comportamientos similares, identificando patrones, asociaciones y otras características relevantes presentes en los datos.

El agrupamiento es una técnica de Machine Learning no supervisado que permite agrupar elementos similares entre sí, con el objetivo de identificar conjuntos con características comunes.

3.1. K-Means

El método de agrupamiento no supervisado busca encontrar la distancia mínima entre un conjunto de datos y el centro de cada grupo, generando así una partición en k grupos a partir de n observaciones. Cada grupo está representado por el promedio de los puntos que lo conforman, y el valor más representativo de cada grupo se llama centroide. El parámetro k , que indica la cantidad de grupos a descubrir, debe establecerse previamente [17].

Una manera de determinar el número de grupos, antes de aplicar el algoritmo K-Means, es mediante el método del codo. Este método calcula la suma de las distancias al cuadrado desde cada punto hasta su centroide asignado en cada iteración de K-Means. Durante cada iteración, se ejecuta el algoritmo con un número distinto de grupos, lo que resulta en un gráfico que muestra la suma de las distancias al cuadrado en función del número de grupos.

Uno de los desafíos principales del algoritmo K-Means es que su resultado puede variar para un mismo conjunto de datos, debido a que los centroides iniciales se seleccionan de forma aleatoria. Esta característica tiene un impacto directo en todo el proceso del algoritmo y puede generar resultados diferentes en cada ejecución.

La Fig. 1 muestra los pasos para la implementación del algoritmo de agrupamiento K-Means.



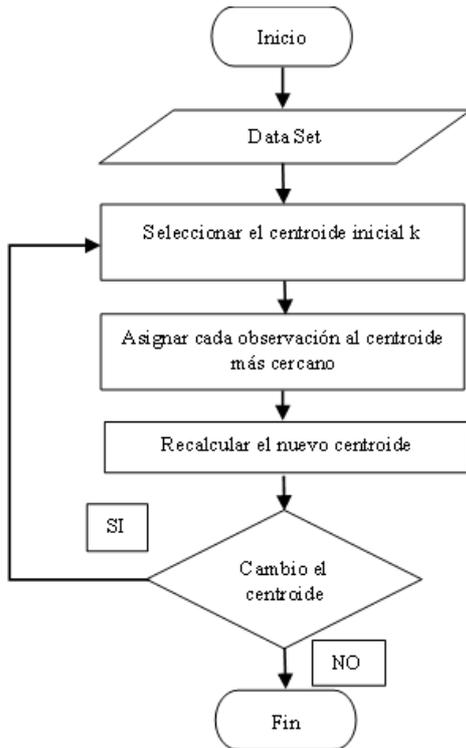


Figura 1: Diagrama de Flujo Algoritmo K-Means

Pasos para implementar el algoritmo K-Means:

- Detallar el número de k grupos que se pretenden encontrar.
- Elegir aleatoriamente k análisis del conglomerado de datos como centroides primarios. $z_1^{(0)}, z_2^{(0)}, \dots, z_k^{(0)}$.
- Estipular los análisis al centroide que se encuentre relativamente más contiguo (usando la distancia euclidiana). $C_i^{(k)}$ indica el vínculo de muestras, cuyo resultado del centroide es $z_j^{(k)}$

$$x \in C_i^{(k)}, \text{ si } d(x, z_i^{(k)}) < d(x, z_j^{(k)}) \quad (1)$$

$$i = 1, 2, \dots, k; i \neq j$$

- Cada uno de los k clústeres debe recalcular los centroides $z_i^{(k+1)}, j = 1, 2, \dots, k$. Entonces, el centroide resultante es:

$$z_i^{(k+1)} = \frac{1}{n_i} \sum_{x \in C_i^{(k)}} x, \quad i = 1, 2, \dots, k \quad (2)$$

Haciendo que n_i sea el número de objetos en $C_i^{(k)}$ Donde: k es en número de grupo. $z_i^{(k+1)}$ centroides primarios. $C_i^{(k)}$ vínculo de muestras.

- Reanudar los pasos tres y cuatro hasta que en las asignaciones no existan variaciones o, por otro lado, se determine el mayor número de iteraciones.

3.2. DBSCAN

El algoritmo DBSCAN es una herramienta diseñada para identificar grupos y ruido en bases de datos espaciales. Define los grupos como el conjunto más extenso de puntos conectados con una densidad específica. Entre sus ventajas se destacan su simplicidad y su capacidad para descubrir agrupaciones con características diversas, revelando valores especiales [18].

Para la implementación de DBSCAN se requiere de forma previa conocer dos parámetros principales que son:

- Epsilon (ϵ): La distancia máxima entre dos puntos cercanos
- MinPts: El número mínimo de puntos cercanos alrededor de un punto especificado, para ser determinado como punto central.

Con los parámetros indicados, cada observación puede ser un punto central, un punto de borde o un punto considerado como ruido.

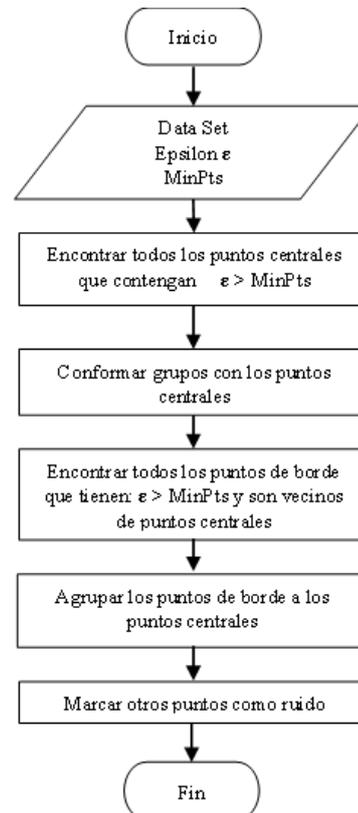


Figura 2: Diagrama de Flujo Algoritmo DBSCAN

El proceso de DBSCAN continúa hasta que todos los objetos han sido procesados. Los puntos que no se asignan a ningún grupo se consideran puntos de ruido, mientras que aquellos que no son ni ruido ni puntos centrales se denominan puntos de borde. De esta manera, DBSCAN construye grupos donde los puntos son clasificados como puntos centrales o puntos de borde, y es posible que un grupo tenga más de un punto central.

El algoritmo comienza seleccionando un punto p arbitrario. Si p es un punto central, se forma un grupo y se incluyen todos los objetos alcanzables desde p . Si p no es un punto central, se visita otro objeto del conjunto de datos. En la Fig. 2 se muestra el diagrama de flujo para la implementación del algoritmo DBSCAN.

4. CASO DE ESTUDIO

4.1. Metodología

En K-Means, es importante considerar los valores atípicos o outliers, ya que el algoritmo se basa únicamente en la distancia entre los elementos y puede verse afectado por la presencia de grupos con tamaños y densidades muy diferentes. Para determinar la cantidad de grupos óptimos, se utiliza el método del codo. Se generan diferentes números de grupos en K-Means y se calcula la similitud entre los elementos dentro de cada agrupación para representarlo gráficamente. El valor en el cual la inercia entre grupos deja de disminuir drásticamente, determina el número óptimo de grupos.

Para representar los nodos en las figuras 3, 4, 5 y 6 se hace uso de los componentes principales 1 y 2 que se obtienen luego de hacer una reducción de la dimensionalidad. El análisis de los componentes principales (PCA) por sus siglas en inglés, es un proceso estadístico que permite reducir la complejidad de los datos a la vez que conserva su información [19].

Para la red de IEEE de 34 nodos se determina que el número de grupos es 3, como se muestra en la Fig. 3. Para la red IEEE de 123 nodos el número de grupos óptimos es de 4 como se indica en la Fig. 4.

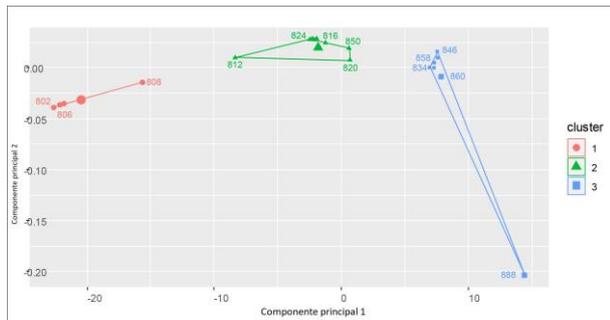


Figura 3: K-Means Clusters IEEE 34 Nodos

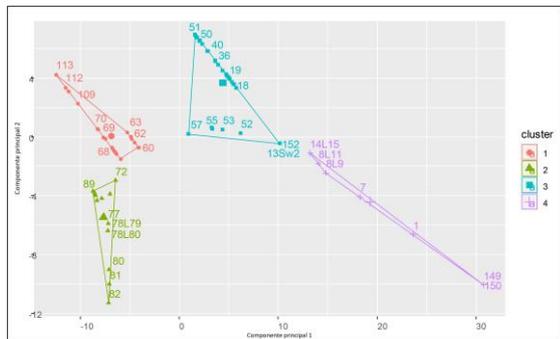


Figura 4: K-Means Clusters IEEE 123 Nodos

El algoritmo DBSCAN establece que, para formar parte de un grupo, un individuo debe tener un número mínimo de vecinos u observaciones dentro de un radio específico, y los grupos deben estar separados por regiones vacías o con pocos individuos. El algoritmo DBSCAN requiere dos parámetros iniciales: Épsilon y Puntos Mínimos. Épsilon representa la distancia máxima entre dos puntos para que sea considerado parte del grupo. En RStudio, se calcula el promedio de las distancias de cada individuo a sus vecinos más cercanos. El valor de Puntos Mínimos (k) es especificado por el usuario, como se muestra en la Tabla 1, y se traza gráficamente en orden ascendente para determinar el punto de inflexión que corresponde al valor de Épsilon.

Tabla 1: Valor de Épsilon y MinPts

Red de Prueba	Número de Observaciones	Épsilon	Puntos Mínimos
34 barras	28	2	3
123 barras	88	3	4

La red IEEE de 34 nodos se conforma de 3 grupos como se visualiza en la Fig. 5, y 3 puntos de ruido descritos en la Tabla 2.

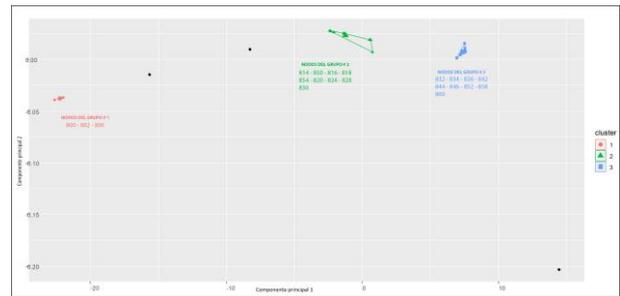


Figura 5: DBSCAN Clusters IEEE 34 Nodos

Tabla 2: Agrupaciones DBSCAN IEEE 34 Nodos

Puntos	Ruido	Clúster		
	0	1	2	3
Borde	3	0	0	0
Seed	0	3	10	12
Total	3	3	10	12

En la red IEEE de 123 nodos se identifican 4 grupos, como se muestra en la Fig. 6. El grupo 0 incluye 7 puntos clasificados como ruidos, según se indica en la Tabla 3.

Tabla 3: Agrupaciones DBSCAN IEEE 123 Nodos

Puntos	Ruido	Clúster			
	0	1	2	3	4
Borde	7	0	0	0	2
Seed	0	5	25	6	43
Total	7	5	25	6	45



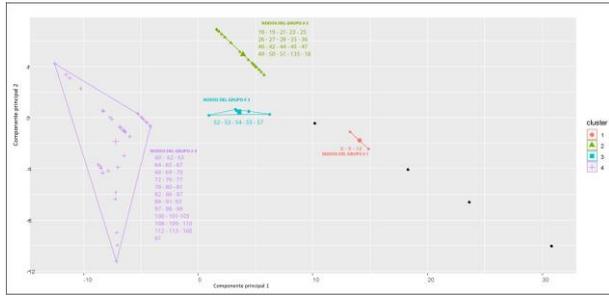


Figura 6: DBSCAN Clusters IEEE 123 Nodos

4.2. Evaluación del Modelo

Para evaluar la precisión de los grupos obtenidos, utilizaremos el coeficiente de silueta, el cual nos indica qué tan similar es un individuo con respecto a los demás individuos dentro del mismo grupo, en comparación con los individuos de los grupos más cercanos. El coeficiente de silueta tiene un rango de valores que va desde negativo hasta 1, y se puede calcular tanto a nivel individual como a nivel de grupo. Un valor cercano a 1 indica que los clústeres son compactos y están adecuadamente separados, un valor de 0 indica que los grupos están superpuestos, y un valor negativo indica que la observación pertenece a otro grupo.

El coeficiente de silueta para la red IEEE de 34 nodos es de 0.81, mientras que para la red de 123 barras es de 0.55, como se muestra en la Fig. 7 y Fig. 8. Estos resultados indican que las agrupaciones obtenidas mediante K-Means son adecuadas, ya que no se observan valores de silueta cercanos a cero o negativos.

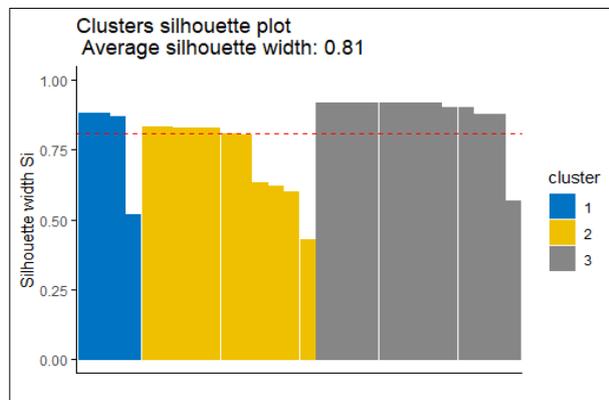


Figura 7: Coeficiente de silueta K-Means red 34 barras

Para calcular el coeficiente de silueta de los grupos obtenidos en DBSCAN, primero identificamos y eliminamos los puntos clasificados como ruido de los datos.

El coeficiente de silueta para la red IEEE de 34 nodos es igual a 0.91 como se indica en la Fig. 9. En el caso de la red IEEE de 123 nodos el coeficiente de silueta es 0.59, resultado que se evidencia en la Fig. 10.

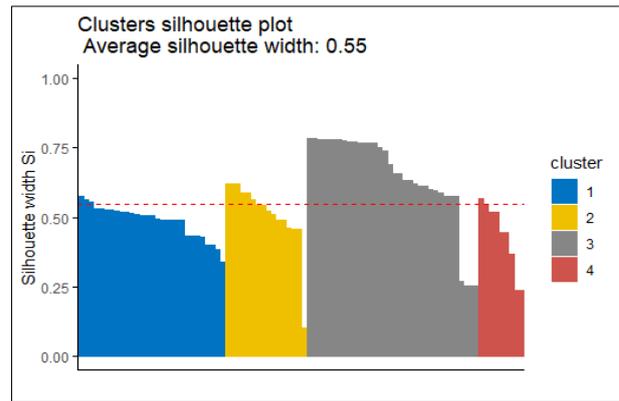


Figura 8: Coeficiente de silueta K-Means red 123 barras

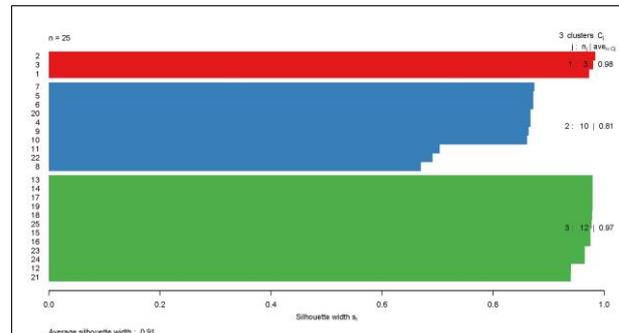


Figura 9: Coeficiente de Silueta DBSCAN Red 34 Nodos

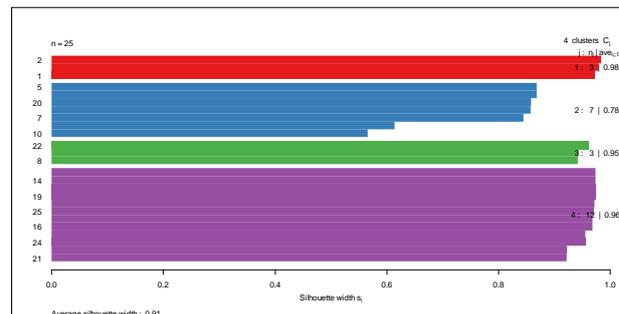


Figura 10: Coeficiente de Silueta DBSCAN Red 123 Nodos

4.3. Implementación del Modelo

Al realizar el análisis, tanto K-Means como DBSCAN en las redes IEEE 34 y 123 nodos, muestran resultados similares, como se observa en la Tabla 4.

Tabla 4: Número de Grupos IEEE 34 y 123 nodos

	K-Means	DBSCAN
Número de grupos IEEE 34 nodos	3	3
Número de grupos IEEE 123 nodos	4	4

En la red IEEE de 34 nodos, se observa una disminución de los niveles de voltaje a medida que aumenta el número de nodos. Los reguladores de voltaje en los sistemas de distribución desempeñan un papel clave en el mantenimiento de los voltajes dentro de los límites establecidos. En el caso de la red de 34 nodos, la fuente se encuentra en el nodo 800 y se produce una

disminución del voltaje hasta el nodo 814, donde se ubica un regulador que contribuye a mantener el voltaje de entrada. Esto podría explicar por qué los grupos están formados por nodos cercanos a la fuente o a un regulador, tal como se visualiza en la Fig. 11.

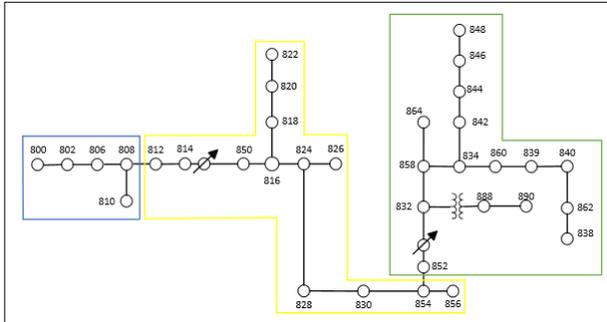


Figura 11: Partición IEEE 34 K-Means

En cada agrupación, los nodos con mayor voltaje corresponden a los nodos anteriores o al nodo donde se encuentra el regulador. Por esta razón, en K-Means, los nodos con valores individuales de coeficiente de silueta más bajos (808, 812 y 888) en cada grupo estarían asociados a nodos con voltajes más altos o más bajos. Sin embargo, en DBSCAN (Fig. 12), los nodos 808, 812 y 888 se consideran ruido.

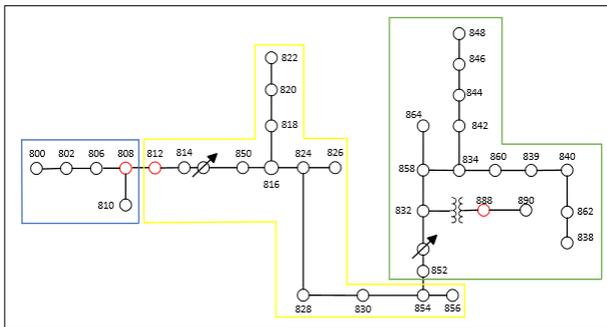


Figura 12: Partición IEEE 34 DBSCAN

Los reguladores de voltaje son responsables de mantener y ajustar el voltaje dentro de rangos seguros tanto en la entrada como en la salida. Si hay una mayor demanda, los reguladores aumentan el voltaje, mientras que, si hay una disminución en la demanda, disminuyen el voltaje. En la red IEEE de 123 nodos, la fuente se encuentra en el nodo 150, seguido de un regulador de voltaje. Por lo tanto, el voltaje en el nodo siguiente, el nodo 149, debe ser ligeramente mayor o igual. La línea principal en la red está compuesta por los nodos 150, 149, 1, 7, 8 y 13, en ese orden, como se muestra en la Fig. 13. Estos nodos deben mantener niveles adecuados de voltaje para garantizar el suministro de energía a las ramificaciones y nodos de la red. Los niveles de variación de voltaje para un funcionamiento seguro deben estar entre el 0.95 y el 1.05 del valor nominal de la fuente.

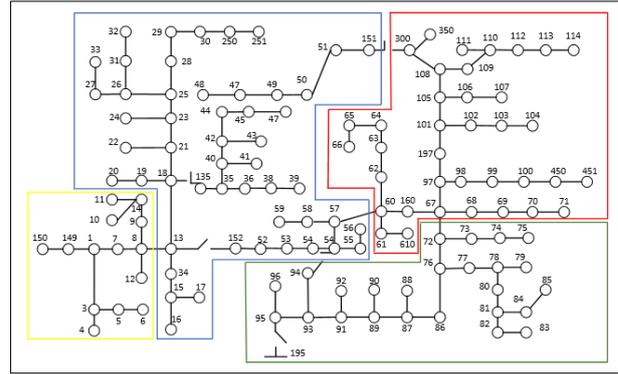


Figura 13: Partición IEEE 123 K-Means

El algoritmo DBSCAN identifica como ruido a los nodos 150, 149, 1 y 7 que pertenecen al grupo 4 donde se encuentra la fuente, y a los nodos 13 y 152, que se encuentran en el grupo adyacente, como se muestra en la Fig. 14. Estas observaciones forman parte de la línea principal. Los tres enlaces de salida del nodo 13 deben permitir el flujo de energía hacia la mayoría de los nodos de la red, por lo tanto, este nodo es muy susceptible a las variaciones (barra piloto). Dependiendo del voltaje en este nodo, podría cambiar variar la configuración del grupo.

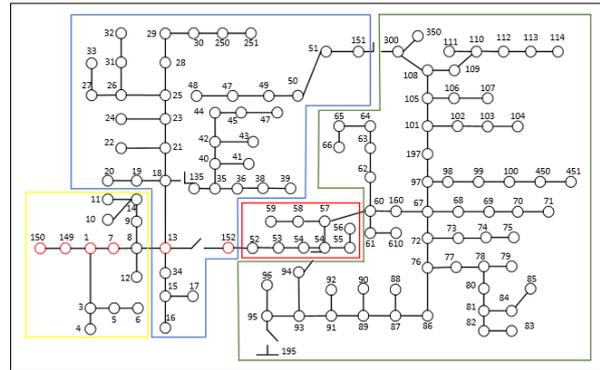


Figura 14: Partición IEEE 123 DBSCAN

5. CONCLUSIONES Y RECOMENDACIONES

Tanto el algoritmo K-Means como DBSCAN mostraron resultados similares en el análisis de la red IEEE de 34 y 123 nodos, lo cual indica que ambas técnicas pueden ser efectivas para agrupar y realizar particiones en redes eléctricas de distribución.

Los nodos de las redes eléctricas de distribución analizadas con valores de coeficiente de silueta bajos sugieren que el nivel de variación de voltaje pueden ser un factor determinante para la conformación del grupo.

Tras comparar los resultados de aplicar las técnicas de minería de datos K-Means y DBSCAN, se observa que ambos algoritmos generan resultados similares al segmentar redes de distribución basadas en los estándares IEEE de 34 y 123 nodos.

Las aplicaciones de los algoritmos K-Means y DBSCAN han demostrado su eficiencia al segmentar la



red de distribución de 34 y 123 nodos, logrando resultados similares a los obtenidos por otros autores mencionados en el estado del arte.

Se recomienda utilizar algoritmos de agrupamiento para analizar y segmentar redes de distribución de diferentes tamaños y características. Esto permitirá obtener una visión más completa y precisa de los grupos presentes en la red.

Se recomienda realizar un análisis de los resultados obtenidos de estos algoritmos. Esto nos ayudará a comprender mejor el funcionamiento de la red y a tomar decisiones más informadas para su óptimo rendimiento.

REFERENCIAS BIBLIOGRÁFICAS

- [1] J. Castro, M. Saad, S. Lefebvre, D. Asber, and L. Lenoir, "Coordinated Voltage Control in Distribution Network with the Presence of DGs and Variable Loads Using Pareto and Fuzzy Logic," *Energies*, vol. 9, no. 2, p. 107, Feb. 2016, doi: 10.3390/en9020107.
- [2] Z. Ernesto Jaramillo, J. R. Castro, T. Castillo, and R. Reategui, "Data Mining in Electrical Distribution Networks: Optimal Location of Pilot Bus," in 2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM), Oct. 2021, pp. 1–5, doi: 10.1109/ETCM53643.2021.9590646.
- [3] E. Velasco-Ramírez, C. Ángeles-Camacho, and M. García-Martínez, "Redes de transmisión inteligente. Beneficios y riesgos," *Ing. Investig. y Tecnol.*, vol. 14, no. 1, pp. 81–88, Jan. 2013, doi: 10.1016/S1405-7743(13)72227-3.
- [4] Y. Chai, L. Guo, C. Wang, Z. Zhao, X. Du, and J. Pan, "Network Partition and Voltage Coordination Control for Distribution Networks With High Penetration of Distributed PV Units," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3396–3407, May 2018, doi: 10.1109/TPWRS.2018.2813400.
- [5] "Community-detection-based approach to distribution network partition," *CSEE J. Power Energy Syst.*, 2022, doi: 10.17775/CSEEJPES.2020.04150.
- [6] B. Zhao, Z. Xu, C. Xu, C. Wang, and F. Lin, "Network Partition-Based Zonal Voltage Control for Distribution Networks With Distributed PV Systems," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4087–4098, Sep. 2018, doi: 10.1109/TSG.2017.2648779.
- [7] Y. Chen, M. G. Fadda, and A. Benigni, "Decentralized Load Estimation for Distribution Systems Using Artificial Neural Networks," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 5, pp. 1333–1342, May 2019, doi: 10.1109/TIM.2018.2890052.
- [8] M. Bahramipناه, M. Nick, R. Cherkaoui, and M. Paolone, "Network clustering for voltage control in active distribution network including energy storage systems," in 2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Feb. 2015, pp. 1–5, doi: 10.1109/ISGT.2015.7131916.
- [9] C. Bonetti, J. Bianchotti, J. Vega, and G. Puccini, "Optimal Segmentation of Electrical Distribution Networks," *IEEE Lat. Am. Trans.*, vol. 19, no. 8, pp. 1375–1382, Aug. 2021, doi: 10.1109/TLA.2021.9475868.
- [10] D. Sharma, K. Thulasiraman, D. Wu, and J. N. Jiang, "A network science-based k-means++ clustering method for power systems network equivalence," *Comput. Soc. Networks*, vol. 6, no. 1, p. 4, Dec. 2019, doi: 10.1186/s40649-019-0064-3.
- [11] Z. M. Ali, A. M. Galal, S. Alkhalaf, and I. Khan, "An Optimized Algorithm for Renewable Energy Forecasting Based on Machine Learning," *Intell. Autom. Soft Comput.*, vol. 35, no. 1, pp. 755–767, 2023, doi: 10.32604/iasc.2023.027568.
- [12] L. R. and others Schneider, Kevin P and Mather, BA and Pal, BC and Ten, C-W and Shirek, Greg J and Zhu, Hao and Fuller, Jason C and Pereira, Jos{\`e} Luiz Rezende and Ochoa, Luis F and de Araujo, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Trans. power Syst.*, vol. 33, no. 3, pp. 3181–3188, 2017.
- [13] R. Dugan, "OpenDSS," EPRI Distribution System Simulator, 2020. <https://sourceforge.net/projects/electricdss/>.
- [14] H. Gu, X. Chu, and Y. Liu, "Partitioning Active Distribution Networks by Using Spectral Clustering," in 2020 IEEE Sustainable Power and Energy Conference (iSPEC), Nov. 2020, pp. 510–515, doi: 10.1109/iSPEC50848.2020.9351132.
- [15] Y. Zou and H. Li, "Study on Power Grid Partition and Attack Strategies Based on Complex Networks," *Front. Phys.*, vol. 9, Jan. 2022, doi: 10.3389/fphy.2021.790218.
- [16] F. Liu, B. Gu, S. Qin, K. Zhang, L. Cui, and G. Xie, "Power grid partition with improved biogeography-based optimization algorithm," *Sustain. Energy Technol. Assessments*, vol. 46, p. 101267, Aug. 2021, doi: 10.1016/j.seta.2021.101267.
- [17] J. MacQueen, "Classification and analysis of multivariate observations," 5th Berkeley Symp. Math. Stat. Probab., pp. 281–297, 1967.

[18] Y. Wang, Y. Li, H. Xie, B. Wu, and Y. Yang, "Cluster division in wind farm through ensemble modelling," *IET Renew. Power Gener.*, vol. 16, no. 7, pp. 1299–1315, May 2022, doi: 10.1049/rpg2.12276.

[19] P. Harrington, *Machine learning in action*. 2012.



investigación están relacionados a la Inteligencia Artificial.

Paúl A. Soto. - Obtuvo el título de Ingeniero en Electrónica y Telecomunicaciones en la Universidad Técnica Particular de Loja en 2009; y de Magister en Gestión de Proyectos de la Universidad Técnica Particular de Loja en el año 2019. Sus campos de



campo de investigación se encuentra relacionado con la Energía y Robótica.

José R. Castro.- Recibió su título de Doctor en Ingeniería Eléctrica (Ingeniería Aplicada) en la ECOLE DE TECHNOLOGIE SUPERIEURE de Montreal Canadá en el 2016. Actualmente, es Docente Investigador de la Universidad Técnica Particular de Loja, y su



Loja (UTPL) en Ecuador. Sus intereses de investigación incluyen minería de texto, minería de datos, aprendizaje automático, aprendizaje profundo con aplicaciones en salud y educación.

Ruth M. Reategui.- obtuvo su grado de Ph.D en la École de technologie supérieure (ÉTS) en Montreal, Canadá en 2019. Actualmente es profesora en el Departamento de Ciencias de la Computación y Electrónica de la Universidad Técnica Particular de



de investigación se encuentra relacionado con el Control, automatización y sistemas inteligentes.

Tuesman D. Castillo.- Recibió su título de Ingeniero en Electrónica y Telecomunicaciones en 2009 de la Universidad Técnica Particular de Loja. Docente a tiempo completo UTPL (2012-). M.A.SC de la École de Technologie Supérieure (ETS) en Montreal (2014-2016), y su campo