

Critical scenarios identification in power system simulations using graph measures and machine learning

Angie Reyes, Yamisleydi Salgueiro, Marco Rivera, Jorge Camargo, Andrés Hernández, and Patrick Wheeler

Abstract—It is essential that electrical power systems are constructed with a reliable and resilient infrastructure. The evaluation of convergence scenarios of the load flow is a technique widely used to study the reliability of energy systems. This paper considers the classification of convergence scenarios under different loading and power generation conditions. Scenarios where the solution is not converging are evaluated using machine learning algorithms. A data set is built from power system topological representation and the simulation of load flows. Algorithms including Support Vector Machine, K-Nearest-Neighbor, and Decision Trees are evaluated and compared. The trained models can be used as a step in the contingency analysis process to be able reduce the computational time and effort in the execution of load flow calculations.

keywords—Critical scenarios identification, Decision trees, K-nearest neighbors, Machine learning, Power system, Support vector machine.

I. INTRODUCCIÓN

LOS eventos de falla en un sistema de potencia pueden provocar interrupciones en el suministro de energía, lo que representa grandes pérdidas económicas para distintos sectores de la sociedad. En la actualidad se busca que las redes de potencia contengan una infraestructura inteligente que sea confiable y robusta [1]. El uso de herramientas computacionales y de optimización es vital, debido al amplio y complejo proceso necesario para estudiar la resiliencia y seguridad de las redes de potencia.

El análisis de contingencias es un proceso que se aplica en el análisis de la seguridad de sistemas de energía eléctrica [2]. Una contingencia representa la interrupción de un elemento específico del sistema de energía. Por ejemplo, una falla

se puede atribuir a condiciones climáticas, fallas técnicas, daños físicos, entre otros. Por lo tanto, el comportamiento de la infraestructura del sistema de potencia se debe simular frente a estos eventos evaluando su impacto en el resto del sistema [3]. El proceso consiste en simular la ocurrencia de cada contingencia y ejecutar el flujo de carga, para luego evaluar los resultados posteriores a la contingencia, y verificar la existencia de violaciones de los límites de operación en distintos elementos de la red. Sin embargo, este método suele ser computacionalmente intensivo y costoso.

Por otro lado, se ha comprobado que la resistencia a estos fallos y eventos se relaciona fuertemente con la topología (estructura) de la red de potencia [4]. Lo que ha motivado el uso de la teoría de grafos en el estudio de este sistema [5]. La estructura topológica de una red de potencia, junto con sus características de comportamiento, como el cumplimiento de los límites operativos de voltaje y el monitoreo de sobrecargas de la red, puede definir la robustez frente a determinadas perturbaciones. En consecuencia, abstraer la complejidad de la red de potencia en un modelo simple de grafo, facilitará su análisis y estudio. A partir del modelo de grafo se pueden extraer medidas de grafo, que permiten obtener una retroalimentación inmediata del diseño de la red, por ejemplo, los componentes críticos del sistema de potencia [6].

Un objetivo del análisis de contingencias es evaluar el comportamiento de la estructura topológica de la red de potencia frente a fallas en distintos escenarios de operación y simulación, mediante el cálculo de flujo de carga (o flujo de potencia) [7], [8]. Cuando el sistema de potencia no encuentran una solución al cálculo del flujo de carga, se indica que el sistema no converge [9]. Evaluar la confiabilidad del sistema significa determinar los componentes más críticos cuando se aplican fallas. Una falla en un componente crítico puede conducir a una inestabilidad del sistema. Por esta razón, es importante verificar el equilibrio de energía entre las cargas y los generadores, y que además los voltajes de bus y los flujos de energía estén dentro de los límites [10]. En un proceso de ejecución de múltiples escenarios, las condiciones iniciales pueden llevar a escenarios de no convergencia y a disminuir la eficiencia del proceso y análisis general. Por otro lado, la optimización en el tiempo de simulación es un papel importante para los diseñadores, operadores y planificadores de sistemas de energía [11].

Entre más escenarios de operación se puedan ejecutar y comprobar, los decisores podrían tomar mejores acciones de diseño y operación de la red basados en el comportamiento

Los autores agradecen el financiamiento del Proyecto de Investigación 1191028 del Programa FONDECYT Regular y FONDAP SERC Chile 15110019.

A. R. es estudiante del Doctorado en Ciencia Aplicada de la Universidad Antonio Nariño, Colombia (e-mail: angreyes@uan.edu.co).

Y. S. trabaja en el Departamento de Ciencias de la Computación en la Facultad de Ingeniería de la Universidad de Talca, Curicó, Chile (e-mail: ysalgueiro@utalca.cl).

M. R. trabaja en el Departamento de Ingeniería Eléctrica en la Facultad de Ingeniería de la Universidad de Talca, Curicó, Chile (e-mail: marcoriv@utalca.cl).

J. C. trabaja en la Facultad de Ingeniería de Sistemas y Computación de la Universidad Antonio Nariño, Colombia (e-mail: jorgecamargo@uan.edu.co).

A. H. trabaja en el Centro de Investigaciones en Ciencias Básicas y Aplicadas (CICBA) de la Universidad Antonio Nariño, Colombia (e-mail: andres.hernandez@uan.edu.co).

P. W. trabaja en University of Nottingham, Nottingham, UK (e-mail: pat.wheeler@nottingham.ac.uk).

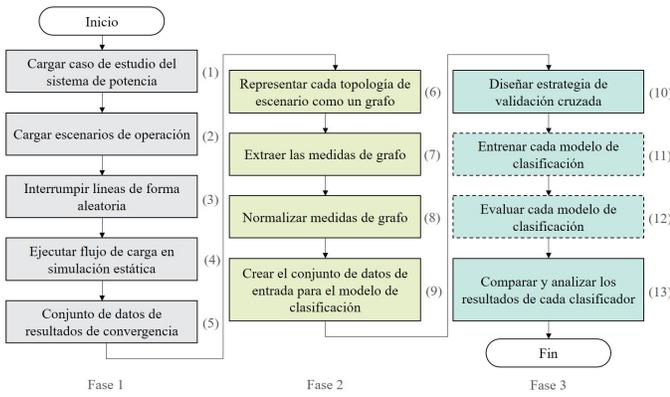


Fig. 1. Diagrama de flujo con la descripción esquemática de las tres fases de la metodología propuesta.

previsto del sistema.

Los métodos de aprendizaje automático (machine learning) se han utilizado en la identificación de fallas [12], predicción de tensión y cargas críticas [13], para estimar el estado de estabilidad o inestabilidad de voltajes en un sistema de potencia a largo plazo [14], y en la predicción de cortes de líneas eléctricas en tiempo real sobre redes en operación [15].

En este artículo se presenta una metodología que inicia con la creación de una base de datos, a partir de la estructura topológica de la red de potencia. Luego se construye un modelo de aprendizaje automático, donde sus características principales se definen a partir de los valores de las medidas de grafo. Se compara el desempeño de tres de diferentes algoritmos de aprendizaje de máquina para el problema de clasificación. También, se analiza la relevancia de las medidas de grafo, mediante las reglas de clasificación de uno de estos clasificadores, con el fin de encontrar relaciones que permitan tomar decisiones sobre la estructura topológica de la red de potencia. Aplicar esta metodología en un estudio de análisis de contingencia, reduciría el tiempo de ejecución de simulaciones, al disminuir la necesidad de ejecutar el cálculo de flujo de potencia en escenarios que no encuentran una solución matemática.

II. METODOLOGÍA

La metodología propuesta en este trabajo considera tres fases principales, como se puede observar en la Fig. 1. La primera fase tiene como objetivo principal crear un conjunto de datos (dataset) mediante la ejecución de simulaciones de contingencia de líneas. La segunda etapa consiste en extraer las características de grafo a partir del conjunto de datos. La última etapa genera como resultado un modelo de aprendizaje de máquina, que, una vez entrenado el modelo, es capaz de predecir los escenarios topológicos en que la red va converger o no.

A. Generación del conjunto de datos mediante simulaciones

La construcción del conjunto de datos se realiza mediante la ejecución de varios flujos de carga. Cada cambio de valor de potencia activa y potencia reactiva en cada uno de los

elementos de generación y de cargas, representa un escenario de simulación.

Además, puede ser el caso en que se establezca o no la aplicación de contingencias $N - 1$ (interrupción aleatoria en una línea de transmisión o transformador) o $N - 2$ (interrupción aleatoria en dos líneas de transmisión o transformadores). Como resultado de la ejecución de cada escenario, se espera saber si el sistema bajo esas condiciones converge o no al cálculo de flujo de carga.

En Python se programa un script que toma como entrada el caso de estudio del sistema (paso 1), y establece un conjunto de escenarios de operación que serán evaluados (paso 2). Por cada escenario de operación se programa una distribución normal para calcular los nuevos valores de potencia activa y potencia reactiva para cada uno de los elementos de tipo carga y generador. A continuación, se determina de forma aleatoria si se aplica o no contingencias de líneas, y de ser así, también de forma aleatoria se desconectan uno o dos elementos de conexión (líneas o interruptores) del sistema de potencia (paso 3). Seguido, se ejecuta el flujo de carga balanceado (paso 4) y se obtiene la información de convergencia del sistema de potencia (paso 5).

Como resultado de la fase 1, se obtiene un conjunto de datos con la información de la estructura topológica de la microrred, los valores de potencia activa y reactiva inicial de los elementos principales, y el valor de convergencia (“true”, si el sistema converge y “false” cuando el sistema no converge). Este conjunto de datos será el insumo de la fase 2, donde se definen las características que conformarán un nuevo conjunto de datos que se utilizará en la construcción del modelo de aprendizaje computacional.

B. Extracción de características a partir de medidas de grafo

La fase 2 de la metodología comienza con el paso 6 (Fig. 1), donde cada escenario se representa mediante un grafo y una etiqueta de convergencia (“converge” o “no converge”). Utilizando como base la teoría de grafos, el sistema de potencia se modela básicamente como un grafo $G = (V, E)$, donde V representa el conjunto de nodos (vértices) y E el conjunto de pares de nodos (aristas). Tomando la estructura física del sistema de potencia, se plantean los elementos de tipo bus, generador y carga como nodos, y los elementos como líneas de transmisión y transformadores, como aristas del grafo. A partir de esta representación simplificada del sistema de potencia, se pueden extraer distintas medidas de red, que constituirán los atributos o características del dataset. A continuación, se describirá cada una de las medidas de grafo utilizadas para construir el vector de características del nuevo conjunto de datos.

1. *Conectividad*: La conectividad es la primera de las medidas que se establece como característica. En este caso, la conectividad retorna como valor un dato de tipo booleano, “true” si el grafo está conectado, o “false” en caso contrario.

2. *Coefficiente de asortatividad*: El coeficiente de asortatividad mide la correlación de las conexiones en el grafo con respecto al grado entre dos pares de nodos conectados. Este coeficiente toma valores entre -1 y 1 , al igual que los coeficientes de correlación.

Los valores cercanos a 1 significan que los nodos conectados tienen grado similar; un valor cercano a -1 significa una correlación entre los nodos conectados con diferente grado; y un valor cercano a 0 significa que no hay una asociación fuerte entre los nodos conectados. Newman [16] definió esta medida como el coeficiente de Pearson de la correlación grado-grado en una red no dirigida. Formalmente, esta medida se define por la Ecuación (1),

$$r = \frac{M^{-1} \sum_i j_i k_i - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}{M^{-1} \sum_i \frac{1}{2}(j_i^2 + k_i^2) - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}, \quad (1)$$

donde j_i y k_i son los grados en ambos extremos de la arista i , y M es el número de aristas.

3. *Índice de Estrada*: El índice de Estrada es una medida topológica de centralidad de nodos del grafo, definido por la Ecuación (2) [17],

$$EE(G) = \sum_{j=1}^n e^{\lambda_j}, \quad (2)$$

donde $\lambda_1, \lambda_2, \dots, \lambda_n$ son los valores propios (Eigenvalues) de su matriz de adyacencia, y n es el conjunto de nodos del grafo.

4. *Centralidad de alcance global*: La centralidad de alcance global de un grafo es el promedio sobre todos los nodos de la diferencia entre la centralidad de alcance local del nodo y la centralidad de alcance local más grande de cualquier nodo en el grafo [18]. Se calcula siguiendo la Ecuación (3). La centralidad de alcance local es una medida de nodo que se calcula como el número de nodos con una distancia dirigida positiva finita desde el nodo i dividido por el número máximo de nodos vecinos.

$$GRC = \frac{\sum_{i \in V} [C_R^{max} - C_R(i)]}{N - 1}, \quad (3)$$

donde N es el número de nodos en el grafo G ; $C_R(i)$ representa la centralidad de alcance local del nodo i , es decir, la proporción de todos los nodos en el grafo que se puede alcanzar desde el nodo i a través de las aristas salientes; y C_R^{max} es la centralidad de alcance local más alto del grafo G .

5. *Cantidad de clanes en un grafo*: Un clan o camarilla (clique) es un conjunto de nodos adyacentes por una arista que los conecta. Se obtiene el tamaño máximo de todos los clanes de un grafo como una medida de grafo. Un clan representa un subgrafo completo del grafo [19].

6. *Cantidad de componentes conectados*: En la teoría de grafos, un componente es un subgrafo en el que pares de vértices están conectados entre sí por caminos, y que no está conectado a vértices adicionales en el resto del grafo.

7. *Conectividad de nodo*: La conectividad de nodo equivale al número mínimo de nodos que al ser eliminados desconectan el grafo. Estos nodos pueden representar un estado crítico, por lo que se relacionan con la robustez del grafo [20].

8. *Conectividad de nodo promedio*: Es el promedio de conectividad de nodo local sobre todos los pares de nodos del grafo G [21]. La conectividad de nodo promedio es obtenida a partir de la Ecuación (4),

$$\kappa(G) = \frac{\sum_{u,v} \kappa_G(u,v)}{\binom{n}{2}}, \quad (4)$$

donde la conectividad $\kappa_G(u,v)$ entre un par de nodos u y v es el número máximo de caminos separados por pares entre ellos, y $\binom{n}{2}$ es el total de pares de nodos del grafo G .

9. *Tamaño del corte*: Un corte es una partición de los nodos de un grafo, de tal forma que se obtengan dos subconjuntos. Por lo tanto, esta medida de grafo es la suma de los pesos de las aristas que cruzan el corte, y esto se conoce como el tamaño del corte.

10. *Expansión de aristas*: Esta medida contiene fuertes propiedades de conectividad. Está dada por la Ecuación (5) [22],

$$i(G) = \text{Min} \frac{e(U, V - U)}{|U|}, \quad (5)$$

donde el mínimo se toma sobre los subconjuntos U de vértices V cuyo tamaño es máximo $\frac{V}{2}$.

11. *Volumen*: El volumen de un conjunto de nodos de un grafo está dado por la Ecuación (6) que representa la suma de los grados de nodos en S [23] y donde d_i es el grado del nodo V_i .

$$\text{vol}(S) = \text{vol}(V(G)) = \sum_{v_i \in S} d_i, \quad (6)$$

12. *Eficiencia global promedio*: Esta medida representa el intercambio de información en toda la red donde la información se intercambia al mismo tiempo. Se calcula utilizando la Ecuación (7) [24],

$$E_{glob}(G) = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{d_{i,j}}. \quad (7)$$

La eficiencia de un par de nodos $E(G)$ en un grafo es el inverso multiplicativo de la distancia de ruta más corta $d_{i,j}$ entre los nodos i y j . La eficiencia global promedio $E_{glob}(G)$ de un grafo es la eficiencia promedio de todos los pares de nodos.

C. Clasificación

En la fase 3 de la metodología se presenta la evaluación y comparación de tres algoritmos de aprendizaje de máquina. Los algoritmos Support Vector Machine (SVM) [25], K-Nearest-Neighbor (KNN) [26], y Decision Trees (DT) [27] se entrenan y se evalúan con la estrategia de validación cruzada, para obtener y comparar su rendimiento en el problema de clasificación de escenarios de convergencia.

La validación cruzada es una estrategia común para calcular el desempeño de los modelos de aprendizaje automático. La validación cruzada se utiliza para estimar los modelos predictivos mediante el particionamiento del conjunto de datos en k

foldeos $n - 1$ foldeos se utilizan como conjunto de entrenamiento y el folde restante se utiliza como conjunto de validación [28].

En la fase 3 de la metodología, se propone la implementación de los pasos 11 y 12 como dependencias del paso 10 (Fig. 1).

Debido a que, dentro de la estrategia de validación cruzada los tres algoritmos de clasificación propuestos, en cada una de las k iteraciones, serán entrenados con el mismo conjunto de datos de entrenamiento, y serán evaluados con el mismo conjunto de datos de validación. El rendimiento de los clasificadores se validará mediante el cálculo del *Accuracy* (porcentaje de exactitud), y se establecerá una comparación y un análisis de los clasificadores (paso 13), con el fin de brindar las condiciones para una correcta toma de decisiones sobre la planificación y operación de un sistema de potencia.

El problema de la clasificación puede ser especificado como un problema para identificar la categoría o clase a la que pertenece una nueva observación basada en un conjunto de datos de entrenamiento que contiene observaciones a las que se les conoce la clase. En este estudio, se examinará, evaluará y comparará el desempeño de los clasificadores SVM, KNN, y DT para la clasificación de escenarios críticos o estables. En este estudio, se usan los valores por defecto de ajuste y parámetros de cada clasificador.

- Support Vector Machine. En problemas de clasificación binaria, SVM depende de encontrar un hiperplano que divida los datos en dos clases diferentes. En SVM los datos de entrada se mapean en un espacio de características de alta dimensión con la ayuda de funciones de kernel [29].
- K-Nearest-Neighbor. KNN es un algoritmo supervisado que clasifica las muestras más similares (por cercanía) basado en el espacio de característica. Cuando se aplica una nueva muestra, el clasificador encuentra el grupo más cercano entre las muestras de entrenamiento [30].
- Decision trees. Los árboles de decisión son un algoritmo que ha sido ampliamente usado para aprender reglas de evaluación en sistemas de energía [31]. Esta técnica principalmente representa una función en cada uno de sus nodos y permite ir construyendo reglas de aprendizaje para la clasificación.

III. SIMULACIONES NUMÉRICAS

El método propuesto se implementó en un sistema de potencia de prueba de dos áreas y cuatro generadores (2A, 4G) [32]. El sistema se presenta en Fig. 2, donde se evidencian dos áreas similares conectadas. Cada una de las áreas tiene dos generadores, el área 1 contiene los generadores $G1$ y $G2$, y en el área 2 se ubican los generadores $G3$ y $G4$. La transferencia de energía del área 1 hacia el área 2 es de $400 MW$. Todas las simulaciones se llevan a cabo en DIGSILENT Power Factory usando el modo Engine con Python [33].

La información de las condiciones de operación del caso base del sistema de potencia 2A4G se presenta en la Tabla I.

En este caso de estudio, se simularon 2016 escenarios operativos del sistema de potencia de prueba para obtener un conjunto de datos significativo. Cada escenario simulado contiene información de la estructura topológica del sistema

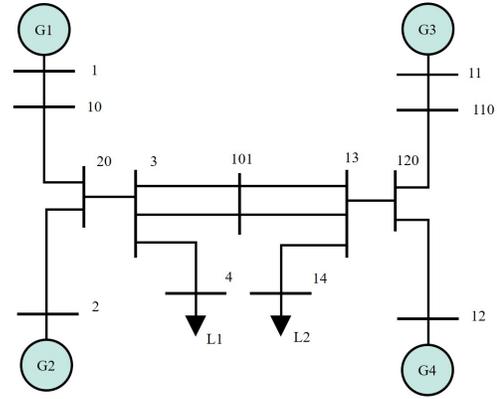


Fig. 2. Diagrama unifilar del sistema de potencia de cuatro máquinas de dos áreas (2A4G).

TABLA I
CONDICIONES INICIALES DE OPERACIÓN DEL SISTEMA 2A4G

Componente	Parámetros
G1	P=700 MW Q=185 MVar
G2	P=700 MW Q=235 MVar
G3	P=719 MW Q=176 MVar
G4	P=700 MW Q=202 MVar
L1	P=696 MW Q=72.01 MVar
L2	P=1272 MW Q=85.32 MVar

de potencia, lo que permite definir un escenario de simulación como una estructura de grafo. Por lo tanto, se obtienen 2016 grafos y un vector de 2016 clases de convergencia (“converge” o “no converge”). Luego, para cada grafo, se extraen las medidas de red topológicas, es decir, no se tendrá en cuenta los pesos de los nodos o vértices; de ser necesario, para calcular una medida se tomará el valor 1 como peso. Cada una de las medidas de red representa una característica del conjunto de datos.

En esta fase se logra obtener una matriz de 2016 muestras con 12 características (medidas de red extraídas de cada grafo), y un vector 2016 etiquetas de clase. Este conjunto de datos se conforma por 1319 (65 %) muestras que hacen referencia a los casos con clase “Converge”, mientras que existen 697 (35 %) muestras con etiquetas de clase “No converge”.

En la Tabla II se presentan los resultados de la ejecución de los tres clasificadores propuestos. Como se mencionó antes, se hace uso de la estrategia de validación cruzada con k -foldeos igual a 5. En las columnas de predicción, junto con las filas correspondientes a cada clasificador, se presenta el resultado de una matriz de confusión seleccionada aleatoriamente de los diferentes grupos validados. Para cada clasificador, las últimas dos columnas resumen la cantidad de predicciones para cada clase y cada fila representa la cantidad de muestras con etiquetas de la clase real “Converge” o “No converge”. La matriz de confusión permite evaluar el desempeño de cada algoritmo de clasificación, en cuanto a la cantidad de aciertos y errores del conjunto de validación para cada una de las clases.

Vale la pena resaltar que, para este conjunto de datos, el algoritmo de clasificación que muestra mejor rendimiento es clasificador KNN.

TABLA II
COMPARACIÓN DE LOS CLASIFICADORES

Tipo de Clasificador	Predicción		
	Actual	No Converge	Converge
SVM	No Converge	103	33
	Converge	5	262
KNN	No Converge	136	0
	Converge	4	263
DT	No Converge	103	35
	Converge	1	266

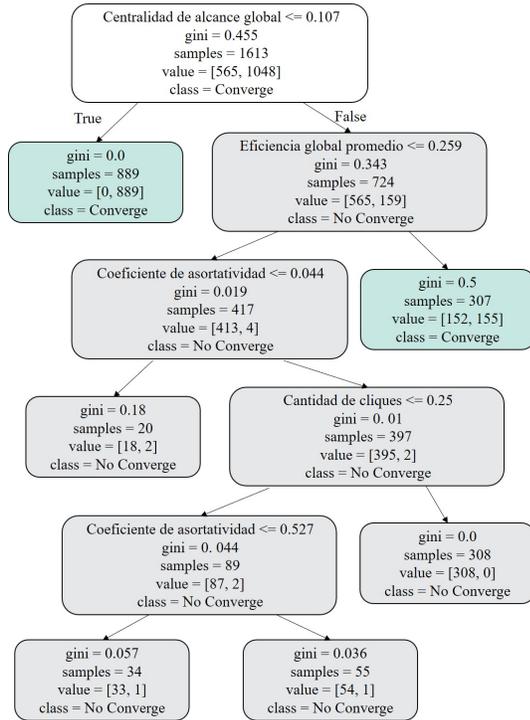


Fig. 3. Árbol de decisión como herramienta gráfica y analítica para analizar las medidas de grafo usadas como características del modelo de clasificación.

IV. DISCUSIÓN

El modelo de árbol de decisión entrenado permite realizar un análisis sobre el comportamiento de las medidas de grafo usadas como características del modelo de clasificación. El grafo del árbol se presenta en Fig. 3. Se puede evidenciar que en el nodo raíz (nodo de color blanco), la medida “Centralidad de alcance global” es la que produce la primera división, es decir, es la característica más relevante en el modelo de clasificación. Esto significa que todo grafo donde esta medida de como resultado un valor menor o igual a 0,107, tiene una gran probabilidad de ser un escenario con convergencia (esto se evidencia con el nodo terminal de color gris) donde el 100 % de los datos que cumplen esta condición pertenecen a la clase “Converge”. Si el valor de la medida “Centralidad de alcance global” es superior a 0,107, es necesario seguir recorriendo los nodos intermedios del árbol.

Se puede evidenciar que, tras la primera división del árbol, se vuelve a dividir el conjunto de muestras con base en la medida del “Índice de Estrada”. Por un lado, los valores que no cumplen la condición de ser menores o iguales a 0,528, se establecen en el nodo terminal de color gris, que

contiene un valor de entropía igual a uno. Por lo tanto, existe la misma frecuencia de clases entre las muestras. De cumplirse la condición, se sigue recorriendo el árbol y así sucesivamente se establecen más niveles del árbol para la clasificación de las muestras con etiqueta “No converge”.

Los nodos categorizados como terminales suelen ubicarse en la parte inferior del árbol y no contienen nodos inferiores a ellos. La función de estos nodos es indicar la etiqueta definitiva de clasificación. Por ejemplo, en el árbol representado en Fig. 3, se puede evidenciar los nodos terminales para la clase “Converge” en color verde. Debido a que en el árbol el primer nodo raíz extiende a un nodo terminal, se puede deducir que la medida topológica más significativa para la clasificación de escenarios que convergen es “Centralidad de alcance global”, además que es más compleja la clasificación de los escenarios que no convergen en el flujo de carga.

V. CONCLUSIONES

En este documento se ha presentado una metodología de clasificación de escenarios de operación de redes de potencia. Los distintos escenarios de operación expuestos a distintas contingencias se simularon en el software DIGSILENT Power Factory y se recopilaron en un conjunto de datos.

La metodología propuesta incluye el entrenamiento y validación del rendimiento de tres algoritmos de clasificación supervisada. Donde se demostró que KNN logra separar con mejor rendimiento, las clases de los escenarios en los que su solución converge o no al flujo de carga.

Los modelos de clasificación presentados en este artículo, pueden ser usados para evaluar la topología de un sistema de potencia, donde se requiera evaluar el diseño topológico frente a desconexiones de elementos físicos. Aplicar previamente este modelo al proceso de simulación de múltiples escenarios de flujos de carga, puede lograr un alto rendimiento de diseño topológico y un bajo de tiempo y esfuerzo computacional.

Los resultados de la clasificación de escenarios vulnerables, permite concluir que las características del escenario basado en medidas de grafos son útiles para la separación de clases. No obstante, la inclusión de pesos en las medidas de grafo y medidas modificadas eléctricamente podrían mejorar los resultados de los clasificadores

Si bien, el caso de estudio presentado se enfoca en la ejecución de simulaciones estáticas, la metodología podría adaptarse para realizar evaluación de sistemas con un conjunto de simulaciones dinámicas, cambiando las variables de entrada del modelo de datos.

Para trabajos futuros, se debe considerar la evaluación de métricas o medidas topológicas en grafos ponderados y dinámicos. También, es necesario aumentar los casos de estudio para evaluar distintas topologías en otros sistemas de potencia.

REFERENCIAS

- [1] A. Öner and A. Abur, “Computationally efficient method for placing reactive power sources against contingencies,” *International Journal of Electrical Power & Energy Systems*, vol. 133, p. 107253, 2021.

- [2] M. Chen, "Contingency re-definition and its application to power system security analysis," in *2011 IEEE/PES Power Systems Conference and Exposition*, 2011, pp. 1–4.
- [3] V. Astapov and I. Palu, "Operation and stability analysis for standalone microgrid through simulations in digsilent software," in *2017 18th International Scientific Conference on Electric Power Engineering (EPE)*, 2017, pp. 1–6.
- [4] J. Fang, C. Su, Z. Chen, H. Sun, and P. Lund, "Power system structural vulnerability assessment based on an improved maximum flow approach," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 777–785, 2018.
- [5] Y. Pei, H. Zhang, X. Gu, and H. Wang, "Research on power grid information model based on artificial intelligence," in *2019 International Conference on Computer Network, Electronic and Automation (ICCNEA)*, 2019, pp. 321–328.
- [6] U. Nakarmi, M. Rahnamay-Naeini, and H. Khamfroush, "Critical component analysis in cascading failures for power grids using community structures in interaction graphs," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 3, pp. 1079–1093, 2020.
- [7] C. Zhai, G. Xiao, H. Zhang, P. Wang, and T.-C. Pan, "Identifying disruptive contingencies for catastrophic cascading failures in power systems," *International Journal of Electrical Power & Energy Systems*, vol. 123, p. 106214, 2020.
- [8] C. Lin, Z. Bie, C. Pan, and S. Liu, "Fast cumulant method for probabilistic power flow considering the nonlinear relationship of wind power generation," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2537–2548, 2020.
- [9] T. Wang, Y. Tang, Y. Huang, X. Chen, S. Zhang, and H. Huang, "Automatic adjustment method of power flow calculation convergence for large-scale power grid based on knowledge experience and deep reinforcement learning," in *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, 2020, pp. 694–699.
- [10] A. Abedi, L. Gaudard, and F. Romerio, "Power flow-based approaches to assess vulnerability, reliability, and contingency of the power systems: The benefits and limitations," *Reliability Engineering & System Safety*, vol. 201, p. 106961, 2020.
- [11] N. Duan, S. Simunovic, A. Dimitrovski, and K. Sun, "Improving the convergence rate of parareal-in-time power system simulation using the krylov subspace," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, 2018, pp. 1–5.
- [12] J. L. Cremer and G. Strbac, "A machine-learning based probabilistic perspective on dynamic security assessment," *International Journal of Electrical Power & Energy Systems*, vol. 128, p. 106571, 2021.
- [13] F. Fachini and B. Fuly, "A comparison of machine learning regression models for critical bus voltage and load mapping with regards to max reactive power in pv buses," *Electric Power Systems Research*, vol. 191, p. 106883, 2021.
- [14] W. M. Villa-Acevedo, J. M. López-Lezama, D. G. Colomé, and J. Cepeda, "Long-term voltage stability monitoring of power system areas using a kernel extreme learning machine approach," *Alexandria Engineering Journal*, 2021.
- [15] J. He and M. X. Cheng, "Machine learning methods for power line outage identification," *The Electricity Journal*, vol. 34, no. 1, p. 106885, 2021, special Issue: Machine Learning Applications To Power System Planning And Operation.
- [16] M. E. Newman, "Assortative mixing in networks," *Physical review letters*, vol. 89, no. 20, p. 208701, 2002.
- [17] E. Estrada, "Characterization of 3d molecular structure," *Chemical Physics Letters*, vol. 319, no. 5-6, pp. 713–718, 2000.
- [18] E. Mones, L. Vicsek, and T. Vicsek, "Hierarchy measure for complex networks," *PLoS one*, vol. 7, no. 3, p. e33799, 2012.
- [19] A. Lancichinetti, S. Fortunato, and J. Kertész, "Detecting the overlapping and hierarchical community structure in complex networks," *New journal of physics*, vol. 11, no. 3, p. 033015, 2009.
- [20] D. R. White and M. Newman, "Fast approximation algorithms for finding node-independent paths in networks," 2001.
- [21] L. W. Beineke, O. R. Oellermann, and R. E. Pippert, "The average connectivity of a graph," *Discrete mathematics*, vol. 252, no. 1-3, pp. 31–45, 2002.
- [22] N. Alon, "On the edge-expansion of graphs," *Combinatorics, Probability and Computing*, vol. 6, no. 2, pp. 145–152, 1997.
- [23] D. Gleich, "Hierarchical directed spectral graph partitioning," *Information Networks*, vol. 443, 2006.
- [24] F. de Pasquale, S. Della Penna, O. Sporns, G. L. Romani, and M. Corbetta, "A Dynamic Core Network and Global Efficiency in the Resting Human Brain," *Cerebral Cortex*, vol. 26, no. 10, pp. 4015–4033, 09 2016.
- [25] D. A. Pisner and D. M. Schnyer, "Chapter 6 - support vector machine," in *Machine Learning*, A. Mechelli and S. Vieira, Eds. Academic Press, 2020, pp. 101–121. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128157398000067>
- [26] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.
- [27] P. H. Swain and H. Hauska, "The decision tree classifier: Design and potential," *IEEE Transactions on Geoscience Electronics*, vol. 15, no. 3, pp. 142–147, 1977.
- [28] C. Schaffer, "Selecting a classification method by cross-validation," *Machine Learning*, vol. 13, no. 1, pp. 135–143, 1993.
- [29] K. Hou, G. Shao, H. Wang, L. Zheng, Q. Zhang, S. Wu, and W. Hu, "Research on practical power system stability analysis algorithm based on modified svm," *Protection and Control of Modern Power Systems*, vol. 3, no. 1, pp. 1–7, 2018.
- [30] D. Bremner, E. Demaine, J. Erickson, J. Iacono, S. Langerman, P. Morin, and G. Toussaint, "Output-sensitive algorithms for computing nearest-neighbour decision boundaries," *Discrete & Computational Geometry*, vol. 33, no. 4, pp. 593–604, 2005.
- [31] Q. Hou, N. Zhang, D. S. Kirschen, E. Du, Y. Cheng, and C. Kang, "Sparse oblique decision tree for power system security rules extraction and embedding," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1605–1615, 2021.
- [32] G. Rogers, "The nature of power system oscillations," in *Power System Oscillations*. Springer, 2000, pp. 7–30.
- [33] DIGSILENT, "Service pack powerfactory 2021 sp3," <https://www.digsilent.de/en/powerfactory.html/>, 2021. [Online]. Available: <https://www.digsilent.de/en/powerfactory.html>