



UNIVERSIDAD NACIONAL
DE COLOMBIA

**Asesoría para implementar dashboard
inteligente para el diagnóstico de redes
eléctricas de nivel de tensión 2, a partir del
análisis de criticidad dado por variables
exógenas y endógenas, y generación de
recomendaciones mediante técnicas de lenguaje
natural**

Código de registro Hermes (UNAL): 54591

Acta de trabajo 1 Contrato CW304777 de 2024 (CHEC) del 3 de septiembre de 2024



28 de octubre de 2024

Índice

1. Introducción	3
2. Estado del Arte	3
3. Objetivos	12
4. Metodología	12
4.1. Extracción, transformación y carga de los datos(ETL)	12
4.2. Desarrollo del sistema de análisis estadístico	14
4.2.1. Pasos en el desarrollo del sistema	14
4.3. Diseño y desarrollo del mockup	15
4.3.1. Pasos en el desarrollo del mockup	15
4.4. Implementación de TabNet como Autoencoder	16
4.4.1. Estructura Matemática de TabNet	16
4.4.2. Estructura del Autoencoder	16
4.4.3. Función de Costo	17
4.4.4. Entrenamiento y Optimización	17
4.5. Desarrollo de Sistema de Recomendación y Consultas Mediante Técnicas de Lenguaje Natural	18
4.5.1. Arquitectura del Sistema	18
5. Resultados	18
5.1. Gráficas y visualización de distribuciones condicionales para eventos en equipos eléctricos	18
5.1.1. Fundamentos del diseño y funcionalidades	18
5.1.2. Resultados de la interfaz de usuario	19
5.2. Mockup del dashboard	20
5.2.1. Pestaña de visualización de mapas	20
5.2.2. Pestaña de chat inteligente	20
5.2.3. Fundamentos del diseño	21
5.3. Análisis de Resultados de TabNet como Autoencoder	21

1. Introducción

En las redes eléctricas de nivel de tensión 2, la criticidad y el análisis de fallos son aspectos fundamentales para garantizar su correcto funcionamiento y la continuidad del servicio. A medida que estos sistemas se vuelven más complejos debido al incremento en la demanda de energía, la interconexión de infraestructuras y la incorporación de nuevas tecnologías, se hace imprescindible contar con herramientas analíticas avanzadas. Estas herramientas no solo deben facilitar el monitoreo en tiempo real de los elementos del sistema, sino también permitir la predicción y prevención de posibles interrupciones, minimizando así el impacto en el servicio y reduciendo los costos asociados a reparaciones no planificadas. En este contexto, los avances en la ciencia de datos y el procesamiento de lenguaje natural (NLP) están ofreciendo soluciones innovadoras que permiten a las empresas optimizar la gestión de riesgos, mejorar las estrategias de mantenimiento preventivo y priorizar intervenciones de manera eficiente basándose en datos concretos y confiables.

El presente proyecto tiene como objetivo principal desarrollar una plataforma integral que permita gestionar y analizar de manera efectiva los datos generados por las redes eléctricas de nivel de tensión 2. Esta plataforma estará diseñada para consolidar y estructurar información crítica, que incluirá mediciones precisas de descargas eléctricas, características constructivas detalladas de los circuitos, históricos de interrupciones del servicio, y variables meteorológicas que podrían influir en el rendimiento de los sistemas. Con este amplio conjunto de datos, se aplicarán técnicas avanzadas de ciencia de datos, tanto descriptivas como predictivas, para identificar patrones clave de fallos, determinar la criticidad de los componentes y anticipar posibles problemas en las redes eléctricas. Estas técnicas permitirán realizar un análisis detallado de la relación entre las condiciones operativas y los fallos, lo que proporcionará una base sólida para optimizar la toma de decisiones a nivel operativo y estratégico.

Además de los análisis descriptivos y predictivos, el proyecto propone el desarrollo de un chatbot inteligente basado en procesamiento de lenguaje natural (NLP), que servirá como asistente virtual para facilitar la consulta de información crítica. Este chatbot no solo será capaz de responder preguntas clave de los usuarios en lenguaje natural, sino que también ofrecerá recomendaciones automáticas sobre las mejores prácticas en la gestión de criticidad y la optimización del rendimiento de las redes eléctricas. La integración de capacidades de análisis geoespacial permitirá que el asistente virtual ofrezca recomendaciones contextualizadas según la ubicación de los equipos, lo que mejorará la precisión de las intervenciones preventivas. Además, el chatbot ofrecerá visualizaciones interactivas para facilitar la interpretación de los datos, ayudando a los operadores y técnicos a tomar decisiones basadas en información comprensible y bien estructurada.

Finalmente, la plataforma incluirá un dashboard inteligente que proporcionará una interfaz visual amigable para la visualización y el análisis de los resultados obtenidos. Este panel de control será capaz de integrar múltiples fuentes de datos y presentarlos de manera clara y coherente, permitiendo a los usuarios monitorear el estado de los sistemas eléctricos, identificar posibles fallos en tiempo real y explorar los análisis predictivos y descriptivos realizados. Además, el dashboard no solo estará disponible en entornos locales para análisis básicos, sino que también se implementará en la nube, lo que permitirá una integración completa con el chatbot y ofrecerá una mayor accesibilidad y escalabilidad. De este modo, los usuarios podrán acceder a la plataforma desde cualquier lugar, realizar análisis más complejos y colaborativos, y beneficiarse de actualizaciones en tiempo real sobre la criticidad del sistema y el estado de los componentes.

Con la integración de estas tecnologías, el proyecto busca transformar el análisis de redes eléctricas en un proceso más eficiente, dinámico y orientado a la prevención. Al incorporar técnicas avanzadas de ciencia de datos y un asistente virtual basado en NLP, la plataforma ofrecerá una solución completa para la gestión de criticidad en redes eléctricas, ayudando a reducir riesgos, mejorar la planificación del mantenimiento y asegurar la continuidad del servicio en un entorno cada vez más exigente.

2. Estado del Arte

En las redes eléctricas el análisis de criticidad tiene como objetivo mejorar la confiabilidad y resiliencia del servicio eléctrico. Debido a la dispersión y diversidad de los componentes en estas redes, se utilizan metodologías especializadas para identificar los elementos críticos y evaluar su comportamiento bajo distintas condiciones operativas y ambientales. Indicadores como SAIDI (Duración Promedio de Interrupción del Sistema), SAIFI (Frecuencia Promedio de Interrupción del Sistema) y CAIDI (Duración Promedio de Interrupción para Clientes Afectados) ofrecen una evaluación cuantitativa de la frecuencia y duración de las interrupciones, lo que facilita tanto la gestión operativa como la planificación estratégica a corto y largo plazo. Estos indicadores ayudan a

desarrollar políticas de mantenimiento ajustadas a las necesidades de la red de distribución eléctrica [?].

En sectores industriales como la minería, se han implementado técnicas de evaluación de criticidad, como Jack-Knife, el Análisis Modal de Fallos y Efectos (AMFE) y matrices de criticidad por riesgo, que optimizan la asignación de recursos y mejoran la productividad [?]. Sin embargo, la aplicación de estas metodologías en redes eléctricas de media tensión presenta desafíos. Aunque el proyecto cuenta con datos precisos sobre fallos, la localización exacta de los elementos afectados es aproximada, lo que dificulta la correcta identificación de los elementos críticos y la priorización de las intervenciones. Esta falta de precisión afecta la planificación del mantenimiento preventivo y correctivo, limitando la efectividad de las acciones. De manera similar, el Árbol de Fallas (FTA) se ha empleado en la industria eléctrica para diagnosticar fallos en transformadores de potencia, descomponiendo los eventos de fallo en sus causas más probables y facilitando la planificación de mantenimientos predictivos. No obstante, el FTA no considera factores exógenos, como las condiciones climáticas o las descargas eléctricas, que pueden aumentar la criticidad de ciertos activos en redes complejas [?]. Esto subraya la necesidad de enfoques que integren la precisión geográfica de los fallos para una gestión más efectiva de la criticidad en redes distribuidas.

Otro enfoque relevante es la automatización del análisis de criticidad mediante herramientas como ETL (Extract, Transform, Load) y PowerBI en circuitos de 13,2 kV. Estas herramientas han demostrado ser eficaces para la integración y visualización de grandes volúmenes de datos, permitiendo a los operadores obtener una imagen clara del estado de la red. Sin embargo, el análisis resultante ha sido mayormente cualitativo, limitando su capacidad para ofrecer diagnósticos basados en métricas cuantitativas. La eficacia de este enfoque depende fuertemente de la calidad y consistencia de los datos recolectados, lo que restringe su aplicabilidad en situaciones que demandan una evaluación numérica rigurosa [?].

La evaluación de criticidad basada en riesgo y severidad ha sido otro método ampliamente utilizado para priorizar las intervenciones de mantenimiento en redes eléctricas. Este enfoque está alineado con los acuerdos de nivel de servicio (SLA) y se enfoca en el impacto operativo de los activos. No obstante, la variabilidad de los activos en redes de nivel de tensión 2 puede complicar la correcta selección y ponderación de los criterios de criticidad, lo que puede derivar en una priorización ineficaz de los recursos de mantenimiento [?]. Esto resalta la importancia de desarrollar una metodología flexible que permita una gestión precisa de los activos, adaptándose a las características específicas de cada red.

Por otro lado, el uso de un índice de confiabilidad ponderado (WI) ha mostrado ser útil para priorizar el mantenimiento de alimentadores de distribución en función de su criticidad. Aunque este enfoque es efectivo, requiere adaptaciones para su aplicación en otros componentes del sistema eléctrico, como transformadores y redes de media tensión [?]. El uso de la medida de betweenness y el algoritmo de adición cíclica ha sido útil para identificar líneas críticas en redes de transmisión, mejorando la redistribución del flujo en situaciones de fallo. Sin embargo, la dependencia de este método en la disponibilidad de datos en tiempo real puede ser un reto para redes sin infraestructura avanzada de monitoreo continuo [?].

En el contexto de tecnologías emergentes, la generación distribuida (DG) y los recursos energéticos renovables han demostrado ser elementos clave para mejorar la confiabilidad de las redes de distribución. Estudios han mostrado que la colocación óptima de DG, utilizando Redes Neuronales Artificiales (ANN) y simulaciones de Monte Carlo, ha logrado reducir en un 40 % el SAIFI y en un 25 % el SAIDI [?]. Este enfoque es particularmente relevante para proyectos que buscan evaluar la criticidad de redes de nivel de tensión 2, donde los indicadores de SAIFI y SAIDI son fundamentales para priorizar el mantenimiento y las intervenciones operativas. No obstante, este enfoque presenta limitaciones cuando no se consideran las reconfiguraciones dinámicas de la red o la interacción con factores exógenos, como las condiciones meteorológicas [?].

La optimización de la ubicación de interruptores automatizados (ASW) mediante algoritmos de optimización estocástica ha demostrado ser eficaz para restaurar rápidamente secciones sanas de la red tras un fallo, reduciendo significativamente los índices de SAIDI y SAIFI. Sin embargo, el éxito de esta técnica depende de la precisión de los datos topológicos y de una infraestructura de comunicación robusta que permita una rápida coordinación entre los sistemas de protección y restauración [?]. De manera similar, la integración de energías renovables y recursos distribuidos mediante simulaciones de Monte Carlo ha mostrado mejoras en la confiabilidad del sistema eléctrico. No obstante, la precisión de los resultados obtenidos con este enfoque depende del número de simulaciones y de las variables consideradas, lo que puede limitar su aplicabilidad en redes complejas o en condiciones meteorológicas adversas [?].

Finalmente, el Método de Monte Carlo (MCS) ha sido ampliamente utilizado para mejorar los indicadores de confiabilidad en redes de distribución. Un enfoque reciente que combinó MCS con simulaciones en Matlab, aplicadas al modelo IEEE 34, logró reducciones del 57.88 % en SAIDI y del 57.63 % en SAIFI, utilizando la se-

lección estratégica de interruptores automáticos y la incorporación de generación distribuida [?]. Otros estudios han investigado la optimización de la ubicación de interruptores de alimentador y el mantenimiento preventivo, logrando mejoras adicionales en SAIDI y SAIFI mediante la instalación de cubiertas aislantes y la gestión de fallos recurrentes [?, ?, ?]. Sin embargo, la efectividad de estas metodologías depende de la precisión de los datos y de las características específicas de cada red, destacando la importancia de combinar mejoras técnicas con una gestión eficiente para garantizar una mejora continua en la confiabilidad de los sistemas eléctricos de distribución.

La integración de la generación distribuida (DG), especialmente con fuentes de energía renovable como paneles solares y turbinas eólicas, ha planteado nuevos retos para la protección de redes de distribución. Los sistemas de protección tradicionales no están diseñados para gestionar flujos de energía bidireccionales ni las variaciones en los niveles de corriente de falla que generan estas fuentes. Esto ha impulsado la adopción de métodos avanzados, como los relés de sobrecorriente direccional (DOCR), las unidades de medición fasorial (PMU) y algoritmos de aprendizaje automático, como las Máquinas de Soporte Vectorial (SVM) [?]. Estos sistemas permiten una detección de fallos más rápida y precisa, mejorando la confiabilidad de las redes con DG. Sin embargo, la implementación de PMU requiere una infraestructura de comunicaciones robusta, lo que puede dificultar su adopción. Además, los métodos de protección diferencial no son siempre efectivos en redes que operan en modos de isla, lo que limita su aplicabilidad en sistemas más complejos. Por lo tanto, se necesita una gestión eficiente de los componentes críticos en redes de media tensión, considerando tanto los atributos constructivos como las variables externas, para optimizar los índices SAIDI y SAIFI y asegurar la continuidad del servicio eléctrico.

Además de los retos en la protección de redes con DG, la restauración eficiente del servicio tras fallos también ha ganado atención, especialmente en sistemas multi-energía. La reconfiguración conjunta de redes eléctricas y térmicas, mediante algoritmos de descomposición distribuida, ha sido una solución para maximizar el uso de recursos energéticos y reducir costos bajo incertidumbre [?]. Aunque este enfoque es efectivo, enfrenta limitaciones similares, como la alta demanda computacional y la falta de enfoques dinámicos para la restauración inmediata. Incluso en redes más convencionales, como los sistemas IEEE 69-bus y 123-bus, la reconfiguración y la optimización de recursos distribuidos de energía (DERs) han demostrado ser eficientes, aunque persisten desafíos cuando hay una alta penetración de DERs o situaciones de contingencia complejas. La precisión de los datos sobre la topología de la red y el estado de los recursos distribuidos sigue siendo un factor crucial para la flexibilidad operativa en estas situaciones [?].

Aunque estas metodologías ofrecen una base sólida, carecen de integración adecuada de variables exógenas críticas, como las condiciones climáticas y las descargas atmosféricas, que pueden influir significativamente en la criticidad de los componentes de las redes eléctricas. Para abordar esta limitación, han surgido enfoques más avanzados en el ámbito de la resiliencia ante desastres naturales. Entre ellos, destacan el uso de sistemas de información geográfica (GIS) y modelos predictivos basados en inteligencia artificial, como aquellos que emplean Redes Neuronales Convolucionales (CNN) y Aprendizaje por Refuerzo, capaces de integrar y procesar grandes volúmenes de datos exógenos y operacionales. Estos métodos no solo permiten una evaluación más precisa bajo condiciones extremas, sino que también optimizan la planificación de medidas preventivas y mejoran la capacidad de respuesta frente a eventos climáticos adversos [?].

Dentro de este marco, la gestión de riesgos en los sistemas energéticos exige un análisis claro de los conceptos de vulnerabilidad y resiliencia. Un marco que distinga entre ambos facilita la identificación de brechas en los datos y modelos actuales, permitiendo una toma de decisiones más informada y un análisis a largo plazo más efectivo [?]. Este enfoque subraya la importancia de mantener la prestación de servicios energéticos incluso frente al fallo de ciertos componentes del sistema. No obstante, lograr este objetivo requiere un desarrollo adicional en la recopilación de datos y la modelación, lo cual puede resultar costoso y depender de políticas públicas. Herramientas como GIS y modelos como INFORM han demostrado su utilidad para proyectar escenarios futuros y evaluar la fragilidad de los sistemas eléctricos, facilitando la implementación de medidas de mitigación ante eventos climáticos extremos [?]. Además, fenómenos como los rayos han sido estudiados mediante sensores satelitales y redes de detección terrestre, lo que ha mejorado la planificación de la protección de las redes frente a descargas atmosféricas [?]. A pesar de ello, estos enfoques deben ajustarse a las particularidades climáticas y operativas de cada región para maximizar su eficacia, como ha sido demostrado mediante el análisis de frontera estocástica (SFA) combinado con GIS, que ha evaluado la eficiencia de las redes de distribución ante fenómenos climáticos, pero requiere adaptación a las características específicas de cada región [?].

De manera similar, un análisis más profundo de la influencia de fenómenos climáticos, como la velocidad del viento, la precipitación y la temperatura en las fallas eléctricas, se ha logrado a través de la evaluación de pro-

piedades multifractales [?]. Sin embargo, este enfoque se limita al estudio de variables climáticas, sin considerar otros factores como las condiciones operativas o las características constructivas del sistema. Para mejorar la precisión, sería necesario integrar estos análisis con datos estructurales y registros históricos de interrupciones, proporcionando un análisis de criticidad más completo. Además, la interdependencia entre el sistema eléctrico y el de distribución de agua plantea nuevos retos en la evaluación de la resiliencia, especialmente bajo condiciones extremas como sequías prolongadas y apagones. La formulación de métricas de resiliencia operativa y estructural, junto con simulaciones de interdependencias, permite cuantificar el impacto de eventos extremos en ambos sistemas [?]. Aunque este enfoque es esencial para garantizar la resiliencia del sistema eléctrico en coordinación con el suministro de agua, su implementación requiere simulaciones extensas, lo que aumenta la complejidad del análisis. Esto subraya la necesidad de desarrollar herramientas más eficientes que integren múltiples variables y aseguren el funcionamiento robusto de infraestructuras críticas bajo condiciones adversas.

En este contexto, un enfoque complementario basado en una matriz de riesgos ha sido utilizado para evaluar el impacto de temperaturas extremas y precipitaciones en diversas fuentes de energía renovable, como la solar, biomasa, biogás e hidroeléctrica [?]. Este método permite identificar riesgos específicos para cada tipo de sistema y ofrece un plan de mitigación estructurado ante eventos climáticos extremos. No obstante, su efectividad depende de la calidad de los datos proyectados, que pueden estar sujetos a incertidumbres significativas. Adicionalmente, los eventos climáticos extremos, como las olas de calor, afectan no solo a la red eléctrica, sino también a la población expuesta a interrupciones en el suministro. Estudios que combinan modelos climáticos avanzados, como el Weather Research and Forecasting (WRF), con modelos de energía en edificios han demostrado que entre el 68 % y el 100 % de la población en ciudades como Atlanta, Detroit y Phoenix podría estar en riesgo de golpe de calor durante eventos simultáneos de olas de calor y apagones [?]. Estos resultados subrayan la vulnerabilidad de infraestructuras críticas en condiciones climáticas extremas, y plantean desafíos adicionales para aquellas regiones con infraestructuras menos robustas, limitando la generalización de los resultados obtenidos en dichas simulaciones.

Siguiendo esta misma línea de análisis, la mejora de la resiliencia en redes eléctricas frente a eventos climáticos severos se ha logrado también mediante el uso del Método de Monte Carlo (MCS), el cual ha permitido cuantificar el impacto en términos de energía no suministrada (ENS) y estimar pérdidas energéticas en distintos escenarios simulados. La distribución Weibull ha sido fundamental para predecir la velocidad del viento, y junto con las curvas de fragilidad, ha facilitado la medición de la resistencia de la infraestructura ante eventos severos [?]. Estos enfoques de optimización, incluyendo el uso de algoritmos genéticos, han mejorado la asignación eficiente de recursos, asegurando que los esfuerzos de endurecimiento se dirijan a las zonas más vulnerables, lo que no solo incrementa la resiliencia sino que también reduce los costos asociados. Además, el MCS ha sido aplicado en otros contextos, como la actividad sísmica, donde ha evaluado la vulnerabilidad de infraestructuras eléctricas críticas como subestaciones y líneas de distribución [?]. A pesar de su efectividad, estos modelos enfrentan retos similares a los planteados en el análisis de fenómenos climáticos, ya que la variabilidad de los datos meteorológicos y la alta demanda computacional limitan su precisión y aplicabilidad, especialmente en redes con menor desarrollo tecnológico o integración de sistemas multi-energía, como la distribución de electricidad y gas natural [?].

En paralelo a estos métodos probabilísticos, se han desarrollado enfoques avanzados basados en aprendizaje por ensamble apilado (stacking ensemble learning) para predecir la duración de los cortes eléctricos causados por tifones, un fenómeno que afecta significativamente la infraestructura eléctrica en muchas regiones. Estos enfoques combinan varios algoritmos potentes como Extra Trees (ET), XGBoost, LightGBM, y otros, permitiendo una predicción precisa gracias a la integración de simulaciones de viento y técnicas avanzadas de preprocesamiento de datos [?]. Aunque este enfoque es muy eficaz, la alta demanda computacional asociada tanto a los modelos basados en Monte Carlo como a los de aprendizaje por ensamble puede ser una limitante en redes eléctricas complejas, como las de media tensión, donde se incrementa significativamente el volumen de datos y el nivel de detalle requerido. Esta problemática también se extiende a las simulaciones probabilísticas utilizadas para evaluar la resiliencia de los sistemas de transmisión, que requieren la inclusión de una gama más amplia de fenómenos climáticos como las lluvias intensas o tormentas eléctricas [?]. Para superar estas limitaciones, innovaciones como el algoritmo basado en la Teoría de Características Estructurales de la Red (NSCT) han mejorado significativamente la identificación de nodos críticos y el tiempo de cómputo, optimizando así la eficiencia operativa de los sistemas de potencia [?].

De manera complementaria, el uso de física informada por aprendizaje automático (PIML) ha demostrado ser una solución avanzada para abordar las complejidades asociadas con la predicción de fallos en redes eléctri-

cas, integrando datos de diversas fuentes, incluidas variables meteorológicas, medioambientales y estructurales. Este enfoque combina modelos predictivos como Gradient Boosted Trees (GBM), que se entrenan con grandes volúmenes de datos, junto con simulaciones físicas detalladas como las proporcionadas por el modelo Weather Research and Forecasting (WRF), que captura datos meteorológicos con alta precisión [?]. Esta combinación no solo permite predecir fallos de manera más precisa, sino también generar recomendaciones estratégicas que optimizan las decisiones de inversión en infraestructura crítica. Las curvas de fragilidad y el análisis de sensibilidad mediante técnicas como dropout loss y perfiles de dependencia parcial (PDPs) aportan una comprensión más profunda de los factores que influyen en las predicciones. Sin embargo, al igual que los enfoques anteriores, este método enfrenta desafíos específicos, como su limitada sensibilidad a eventos de menor intensidad y la falta de evaluación en contextos de tormentas más complejas, lo que sugiere la necesidad de seguir desarrollando y ajustando estos modelos para escenarios meteorológicos más diversos.

Por otro lado, el modelado de redes eléctricas como sistemas complejos interconectados, que dependen de múltiples fuentes de energía, requiere una perspectiva integral para comprender su comportamiento dinámico. Para abordar esta complejidad, se han empleado modelos basados en la teoría de redes y sistemas ciberfísicos, complementados con simulaciones que utilizan métricas de centralidad como el net-ability y el grado entrópico. Estas métricas permiten identificar puntos vulnerables en redes con generación distribuida, evaluando la importancia de los nodos dentro de la red [?]. Sin embargo, aunque estas herramientas proporcionan información valiosa sobre la estructura de la red, no logran capturar de manera adecuada el comportamiento dinámico ni las interacciones entre los componentes durante perturbaciones. Esto resalta la necesidad de desarrollar métodos más precisos que incluyan aspectos dinámicos y temporales para evaluar con mayor exactitud la resiliencia del sistema.

En cuanto a la cuantificación de la resiliencia, se han desarrollado métodos que permiten analizar el rendimiento de los sistemas energéticos frente a eventos de alto impacto y baja probabilidad (HILP), así como de bajo impacto y alta probabilidad (HPLI). Este marco proporciona una visión clara sobre cómo integrar la resiliencia en la planificación de infraestructuras críticas, permitiendo diferenciar entre resiliencia y fiabilidad. Dichos métodos permiten cuantificar mejoras en resiliencia a través de simulaciones detalladas, pero su implementación sigue siendo compleja y costosa, dado que requiere altos niveles de coordinación y gobernanza entre los diferentes actores involucrados en la operación de estas infraestructuras [?]. No obstante, la posibilidad de integrar resiliencia en las decisiones de planificación a largo plazo resulta fundamental para garantizar la seguridad y eficiencia de los sistemas eléctricos en situaciones de riesgo.

En complemento con este enfoque, el fenómeno de los fallos en cascada ha sido un área de investigación clave, particularmente por su capacidad de desencadenar fallos sistémicos en las redes de energía. Este tipo de análisis se basa en la simulación de fallos en cascada utilizando modelos de corriente alterna (AC), junto con algoritmos de optimización metaheurística que permiten identificar y clasificar las ramas críticas dentro de la red según su vulnerabilidad [?]. Al identificar no solo las ramas críticas, sino también clasificarlas en función de su impacto potencial, este enfoque facilita la planificación de la respuesta a fallos tanto a corto como a largo plazo. Sin embargo, las simulaciones de fallos en cascada requieren un elevado poder computacional, lo que plantea desafíos importantes para su implementación práctica en redes de gran tamaño y complejidad.

En este sentido, la creación de un sistema de soporte a la decisión (RV-DSS) ha permitido integrar las interdependencias entre diferentes infraestructuras críticas, mejorando la capacidad de recuperación de las redes al coordinar los procesos de recuperación entre múltiples sistemas interconectados. Un caso de estudio en North Argyll, Escocia, demostró que este sistema puede mejorar significativamente la capacidad de respuesta ante fallos y optimizar la planificación adaptativa [?]. No obstante, su implementación en redes más grandes requiere un alto nivel de coordinación entre múltiples propietarios y operadores de infraestructuras, lo que puede complicar su adopción debido a los costos elevados y a la complejidad organizativa involucrada.

Complementando estos avances, las Redes Neuronales Bayesianas (BNN) han sido empleadas para predecir fallas en sistemas de distribución, integrando datos históricos y meteorológicos para generar intervalos de confianza útiles en la toma de decisiones en tiempo real. Sin embargo, la precisión de estos modelos se ve limitada por la falta de integración de datos constructivos, como los relacionados con apoyos eléctricos, transformadores, switches y redes de media tensión, lo que es esencial para un análisis de criticidad más robusto [?]. En el contexto de tu proyecto, la incorporación de estos datos estructurales podría mejorar significativamente la representación física del sistema eléctrico, optimizando tanto la identificación de componentes críticos como las estrategias de mantenimiento.

En la misma línea, el uso de aprendizaje profundo de refuerzo multi-agente (MADRL) ha mostrado mejo-

ras significativas en la coordinación y programación de sistemas de almacenamiento de energía móvil para la restauración de carga. Este enfoque se basa en la utilización de una red Multi-Agent Double Deep Q-Network (P-MADDQN), permitiendo que múltiples agentes interactúen para optimizar la distribución de recursos en función de la criticidad de las cargas esenciales y no esenciales en redes eléctricas. Sin embargo, aunque este modelo se adapta bien a sistemas que requieren decisiones en tiempo real, su aplicación puede enfrentar barreras debido a la necesidad de una infraestructura de comunicación robusta, lo que resulta costoso y difícil de implementar en redes eléctricas que no operan con datos en tiempo real. Además, la falta de un conjunto de datos completo que represente el comportamiento de la carga y las interacciones entre los componentes de la red podría limitar la efectividad del modelo en ciertos contextos [?].

De manera similar, la optimización y el análisis predictivo en redes eléctricas se han enfocado principalmente en mejorar la resiliencia operativa y mitigar fallas, especialmente ante condiciones climáticas adversas, utilizando métricas de riesgo y herramientas avanzadas para priorizar el mantenimiento de activos críticos. Estas estrategias han demostrado ser altamente efectivas en la identificación de riesgos meteorológicos, pero tienden a pasar por alto otros factores importantes, como las características constructivas y operativas del sistema eléctrico, las cuales son esenciales para un análisis integral [?]. En este contexto, este sesgo hacia fenómenos meteorológicos limita la capacidad de los modelos para proporcionar una evaluación completa, lo que sugiere la necesidad de integrar una visión más amplia que incluya factores estructurales y operativos dentro de los análisis de criticidad. Por otro lado, un área de creciente interés en la optimización del rendimiento de las redes eléctricas es la detección de fallos en sistemas fotovoltaicos a gran escala. Este proceso se ha perfeccionado mediante el uso de técnicas de análisis de residuos y detección de puntos de cambio (CP), lo que permite identificar fallos con una precisión del 97.3 % en series temporales de rendimiento. Estas técnicas son cruciales para diferenciar entre fallos irreversibles, como la degradación de los módulos fotovoltaicos, y fallos reversibles, como el soiling (acumulación de suciedad), mejorando significativamente la capacidad de respuesta y el mantenimiento en tiempo real de estos sistemas energéticos. No obstante, la efectividad de estas técnicas depende en gran medida de la disponibilidad de datos constructivos del sistema fotovoltaico, lo que podría limitar su aplicabilidad a instalaciones con información limitada o infraestructuras más simples [?]. A pesar de esto, esta metodología representa un avance clave en la eficiencia operativa de los sistemas fotovoltaicos y podría aplicarse en redes eléctricas híbridas, donde la detección temprana de fallos es crítica para optimizar el rendimiento y garantizar la continuidad operativa. En línea con esta tendencia, el enfoque híbrido basado en EnsembleBoost ha demostrado ser una solución prometedora para la detección de fallos en tiempo real en redes eléctricas. Este método combina técnicas de ensemble learning, como el Gradient Boosting Decision Trees (GBDT), con comunicaciones inalámbricas para mejorar la precisión y reducir la latencia en la detección de fallos, alcanzando un 95.60 % de precisión y un recall del 91.90 % en instancias fallidas [?]. Además, este enfoque destaca no solo por su alto rendimiento, sino también por su capacidad para reducir en un 35 % la latencia, superando a métodos tradicionales como Random Forests y Decision Trees. Complementando este enfoque, la detección automática de fallos en motores de inducción de tres fases mediante imágenes térmicas y técnicas avanzadas como Scale-Invariant Feature Transform (SIFT) y la red neuronal convolucional AlexNet ha demostrado una precisión del 100 % en la detección de fallos [?]. Aunque esta técnica ofrece una solución no invasiva para el monitoreo de motores, su implementación depende de la disponibilidad de imágenes térmicas, lo que puede no estar garantizado en todos los sistemas actuales. En conjunto, la integración de estos enfoques refleja la importancia de las tecnologías avanzadas en la mejora de la detección de fallos y la operatividad de las redes eléctricas.

Dado que un análisis detallado y preciso de las redes eléctricas requiere la integración de diversas fuentes de información, como datos climáticos, operativos y estructurales, es necesario implementar estrategias adecuadas de fusión de datos. Este proceso permite consolidar la información de manera coherente, facilitando el análisis de grandes volúmenes de datos heterogéneos. La fusión de datos no solo simplifica la identificación y diagnóstico de fallos, sino que también permite una evaluación más precisa de la resiliencia del sistema. Al estructurar los datos de forma uniforme, se optimizan los modelos predictivos y analíticos, lo que mejora la toma de decisiones para el mantenimiento y operación de las redes eléctricas, especialmente en contextos de alta variabilidad e incertidumbre[?].

De manera complementaria, la fusión a nivel de datos consiste en la integración directa de datos sin procesar provenientes de múltiples fuentes, antes de realizar cualquier tipo de procesamiento o extracción de características. Esta técnica ha sido fundamental para redes eléctricas de nivel de tensión 2, ya que permite fusionar información de sensores, condiciones meteorológicas y registros de fallos, mejorando la detección y predicción de eventos críticos. En la teledetección oceánica, por ejemplo, se ha utilizado el Formalismo Multifractal Micro-

canónico para fusionar datos de diversas fuentes y caracterizar fenómenos como la turbulencia oceánica [?]. Este formalismo permite identificar puntos singulares que contienen la información más relevante en los datos, algo que podría ser útil para identificar puntos críticos en la infraestructura eléctrica. No obstante, la implementación de esta técnica en redes eléctricas requiere de datos de alta resolución, lo que puede aumentar la complejidad y los costos computacionales, representando un desafío en términos de infraestructura de datos y calidad.

En paralelo, se ha introducido el concepto de fusión de datos en el borde de la red, que permite procesar información de dispositivos IoT en tiempo real sin depender de la nube [?]. Esta técnica utiliza algoritmos de inteligencia artificial desplegados en dispositivos periféricos, lo que reduce la latencia y el uso de ancho de banda, mejorando así la eficiencia operativa en redes eléctricas. Esto es particularmente útil en redes de nivel de tensión 2, donde las condiciones en tiempo real requieren respuestas inmediatas. Sin embargo, uno de los desafíos es asegurar que los algoritmos de fusión sean lo suficientemente robustos para operar en entornos heterogéneos y con datos de calidad variable, lo que puede limitar la efectividad de este enfoque en ciertos contextos.

Un enfoque relacionado es la fusión multimodal, que combina datos de diversas modalidades, como sensores, condiciones climáticas y registros operacionales. Técnicas de fusión temprana, tardía y jerárquica han sido desarrolladas para optimizar la integración de estas fuentes de datos en redes eléctricas [?]. En el contexto de redes de nivel de tensión 2, esta fusión de datos puede mejorar significativamente la toma de decisiones al combinar información relevante de diferentes modalidades, mejorando la eficiencia y resiliencia de las redes. Sin embargo, un desafío importante es la heterogeneidad de las fuentes de datos y la necesidad de realizar un preprocesamiento para unificar los formatos y escalas, lo que requiere una inversión considerable en infraestructura. Además, en sistemas estructurales, se han aplicado técnicas de fusión de datos para combinar métodos de prueba como espectroscopía de terahercios y termografía infrarroja para mejorar la relación señal-ruido y optimizar la detección de fallos [?]. Este enfoque podría ser adaptado para redes eléctricas, mejorando la detección de anomalías en componentes críticos mediante la combinación de diversas pruebas y mediciones.

Por otro lado, el desarrollo de un modelo de fusión multicaracterística basado en el algoritmo de apilamiento ha permitido combinar regresión de soporte vectorial (SVR) y redes neuronales de memoria a largo plazo (LSTM) para estimar el estado de salud de baterías [?]. En este modelo, la SVR se utiliza para construir múltiples modelos a partir de diferentes características (como temperatura, voltaje y corriente), mientras que la LSTM fusiona los resultados para mejorar la precisión de las predicciones. Este enfoque puede ser particularmente útil en redes de nivel de tensión 2, donde la combinación de variables críticas como condiciones meteorológicas y descargas eléctricas puede mejorar la predicción de fallos. Sin embargo, el desafío radica en seleccionar adecuadamente las características relevantes, ya que una mala elección puede generar redundancia y disminuir la precisión del modelo. Además, se requiere una infraestructura computacional robusta para manejar el procesamiento de grandes volúmenes de datos.

El uso de modelos integrados basados en la fusión de características multisensoriales y el aprendizaje profundo ha mostrado gran potencial para monitorear redes eléctricas de nivel de tensión 2. Un ejemplo de esto es el uso de una red neuronal convolucional paralela (PCNN) para extraer características de múltiples sensores, seguida de redes neuronales residuales (DRN) y LSTM bidireccional para procesar las características fusionadas, mejorando la identificación de fallos en tiempo real [?]. Este modelo combina las fortalezas de las CNN para aprender características espaciales con las LSTM para capturar relaciones temporales. No obstante, su implementación presenta desafíos, ya que requiere grandes cantidades de datos para entrenar las redes neuronales profundas, lo que implica una inversión considerable en infraestructura y recursos computacionales.

Finalmente, la fusión a nivel de decisión permite integrar los resultados de diferentes modelos o sensores para generar decisiones finales más robustas. En el diagnóstico de fallos y monitoreo de estado de equipos eléctricos, esta técnica ha demostrado mejorar la confiabilidad del sistema, especialmente al manejar información contradictoria o parcial. Un enfoque basado en la teoría de evidencia de Dempster-Shafer y la entropía de Deng ha sido utilizado para diagnosticar fallos en convertidores de energía eólica, gestionando de manera efectiva la incertidumbre en las señales de fallo [?]. Este enfoque, junto con el uso de la divergencia de Jensen-Shannon para medir conflictos entre evidencias de diferentes sensores, mejora el monitoreo de equipos críticos en redes eléctricas [?]. Sin embargo, la implementación de estos sistemas requiere de una calibración precisa y una infraestructura robusta para procesar grandes volúmenes de datos en tiempo real.

Además, se han desarrollado estrategias adaptativas de fusión de decisiones que mejoran el diagnóstico de fallos mediante señales de múltiples canales y modelos basados en redes neuronales convolucionales [?]. Estas metodologías podrían aplicarse en redes eléctricas de nivel de tensión 2 para analizar condiciones operativas de componentes críticos, permitiendo la identificación temprana de comportamientos anómalos y la anticipación de posibles fallos. Sin embargo, la implementación de estos enfoques requiere conjuntos de datos amplios y re-

presentativos para entrenar los modelos de manera eficaz, así como una considerable capacidad computacional para garantizar una detección rápida y precisa.

El siguiente cuadro presenta un resumen técnico y detallado del estado del arte en la evaluación y mejora de la criticidad y resiliencia en redes eléctricas. Se abordan las principales áreas de investigación actuales evaluación de criticidad de activos, integración de variables climáticas y estrategias avanzadas de fusión de datos detallando las metodologías aplicadas, las contribuciones principales, las limitaciones técnicas identificadas, las referencias relevantes y los desafíos comunes asociados a cada enfoque.

Área de Investigación	Enfoque Específico	Metodologías Aplicadas	Principales Contribuciones	Limitaciones Técnicas Identificadas	Referencias	Desafíos a nivel del proyecto
Evaluación de Criticidad de Activos	Enfoques Tradicionales	Indicadores de Fiabilidad (SAIDI, SAIFI, CAIDI), Jack-Knife, AMFE	Cuantificación de interrupciones y evaluación de criticidad de activos	Falta de granularidad; localización imprecisa de fallos; no consideran variables exógenas	[?], [?]	La falta de integración de variables exógenas, como clima y descargas atmosféricas, limita la evaluación de criticidad
	Automatización Basada en Datos	Procesos ETL, Herramientas BI (PowerBI)	Integración y visualización de datos operacionales a gran escala	Análisis cualitativo; ausencia de métricas cuantitativas robustas; datos heterogéneos	[?]	
	Enfoques Emergentes	Generación Distribuida (DG), Simulaciones de Monte Carlo (MCS)	Optimización de la ubicación de DG y mejora de la confiabilidad mediante simulaciones	No consideran reconfiguraciones dinámicas; interacción limitada con condiciones climáticas; alta demanda computacional	[?], [?]	
Integración de Variables Climáticas	Factores Externos y Ambientales	Sistemas GIS, Sensores Satelitales, Modelos de IA (CNN, Aprendizaje por Refuerzo)	Proyecciones climáticas; planificación preventiva ante descargas; procesamiento de datos exógenos	Dependencia de datos geográficos precisos; acceso limitado a sensores; infraestructura necesaria para IA	[?], [?], [?]	La incorporación de datos exógenos, aunque ya disponibles, exige técnicas avanzadas para manejar grandes volúmenes de información
	Modelos Predictivos Avanzados	CNN, MCS, Curvas de Fragilidad	Planificación preventiva; evaluación de resistencia; predicción de eventos extremos	Requieren grandes volúmenes de datos; alta demanda computacional; necesidad de datos estructurales precisos	[?], [?]	
Estrategias Avanzadas de Fusión de Datos	Fusión de Datos Directa	Formalismo Multifractal Microcanónico, Edge Computing	Identificación de puntos críticos; procesamiento en tiempo real desde dispositivos IoT	Necesidad de datos de alta resolución; complejidad y costos computacionales; variabilidad en la calidad de los datos	[?], [?]	Requieren desarrollo de modelos eficientes para integrar datos heterogéneos y mejorar la precisión en la evaluación de criticidad
	Fusión Multimodal y de Características	Integración Multimodal, Pruebas Estructurales, Modelos Multi-característica	Integración de datos diversos; detección de fallos; predicción del estado de salud de baterías	Preprocesamiento complejo; pruebas costosas; necesidad de infraestructura robusta	[?], [?], [?]	
	Fusión Basada en Decisiones	Fusión Multisensorial con CNN/LSTM, Teoría de Dempster-Shafer, Fusión Adaptativa	Detección de fallos en tiempo real; manejo de incertidumbres; mejora en diagnóstico de fallos	Requiere grandes cantidades de datos; limitaciones en infraestructura; precisión dependiente de la calidad de datos	[?], [?], [?]	

Tabla 1: **Resumen del Estado del Arte: Técnicas y Limitaciones en la Evaluación y Mejora de la Criticidad y Resiliencia en Redes Eléctricas**

3. Objetivos

- Construir una base de datos estructurada que incluya datos de mediciones de descargas eléctricas, características constructivas de los circuitos, históricos de interrupciones, variables meteorológicas y documentos de normatividad de interés, con el fin de realizar un análisis de criticidad en redes eléctricas de nivel de tensión 2.
- Implementar técnicas de ciencia de datos que permitan identificar patrones de criticidad a partir de la base de datos construida, abarcando análisis descriptivo y predictivo, junto con visualizaciones interactivas para facilitar la interpretación y toma de decisiones.
- Desarrollar una estrategia de procesamiento de lenguaje natural, como un chatbot, para la consulta de patrones de interés y la generación de recomendaciones orientadas al monitoreo de niveles de criticidad en redes eléctricas de nivel de tensión 2.
- Construir un dashboard inteligente que integre los modelos de ciencia de datos y de procesamiento de lenguaje natural desarrollados, tanto en un ambiente local (para análisis básico) como en un entorno de nube (para análisis completo con la integración del chatbot).

Estos objetivos buscan asegurar una implementación efectiva y estratégica de soluciones analíticas que faciliten el monitoreo, análisis y gestión de la criticidad en las redes eléctricas, alineándose con las necesidades operativas y de toma de decisiones de la empresa.

4. Metodología

4.1. Extracción, transformación y carga de los datos(ETL)

Un proceso ETL resulta fundamental para asegurar la integración y calidad de los datos en el proyecto de criticidad del sistema de red de media tensión de CHEC. Su importancia radica en reunir datos, transformarlos a un formato consistente y adecuado para el análisis, y hacerlos disponibles para los modelos de IA que se desarrollarán en etapas posteriores. Este proceso garantiza que los datos sean precisos, completos y preparados para un análisis eficiente, permitiendo decisiones informadas y efectivas. A continuación, se detallan cada uno de los pasos.

La consolidación de datos y eliminación de columnas irrelevantes son pasos esenciales para preparar los datos para su análisis y modelado. El código propuesto procesa archivos CSV organizados por prefijos específicos, como **APOYOS**, **EVENTOS**, **REDMT**, **SWITCHES**, y **TRAFOS**. En cada categoría, se eliminan columnas redundantes o innecesarias antes de consolidar los archivos, estructurando los datos de forma uniforme y eficiente para enfocar el análisis en las variables más relevantes y reducir el tamaño de los datos para facilitar su manejo. El proceso inicia con la verificación de archivos por prefijo y unificación de aquellos con estructuras de columnas coincidentes, dada la información distribuida en distintas tablas. En caso de archivos únicos para un prefijo, el script elimina duplicados y descarta columnas irrelevantes. Para múltiples archivos, evalúa la coincidencia de columnas y consolida solo aquellos con estructuras uniformes, asegurando así la integridad de los datos y reduciendo discrepancias de columnas que pueden surgir al trabajar con varias versiones o fuentes. El conjunto de datos se guarda en una carpeta independiente de los datos originales, preservando estos últimos para respaldo o ajustes en el proceso. Esta organización asegura accesibilidad al conjunto de datos final y facilita su análisis. Al reducir las columnas a las necesarias y eliminar duplicados, el archivo resultante se optimiza en términos de almacenamiento y procesamiento, lo cual es clave para compatibilidad en análisis avanzados e integración de modelos de IA en el sistema de red de media tensión de CHEC.

La primera tabla procesada es **EVENTOS**, que contiene interrupciones desde 2019 hasta junio de 2024. Este conjunto de datos se somete a una serie de transformaciones para limpiar la información, garantizando que los datos sean consistentes y adecuados para el análisis de criticidad. Se seleccionan solo registros donde **tipo_causa** sea No programada, eliminando luego esta columna por su redundancia dado que permanece constante para todos los registros. Un segundo filtro en **tipo_elemento** mantiene valores de tensión menores o iguales de

33KV, correspondientes a fallos en la red de media tensión de CHEC. Esto permite enfocar el análisis en eventos específicos y elimina datos no pertinentes para los objetivos del estudio. En la columna **causa**, un mapeo con un diccionario reemplaza códigos numéricos por descripciones, como “Apertura por ejecución de esquemas suplementarios”, “Falla en redes de baja tensión”, facilitando la interpretación de datos. Finalmente, las columnas **inicio** y **fin** se convierten al formato de fecha y hora para cálculos temporales, como la duración de eventos y patrones de ocurrencia. Estas transformaciones resultan en una estructura más coherente, facilitando su análisis.

La tabla **APOYOS** incluye información sobre la infraestructura de apoyo en la red de media tensión y que se somete a varias transformaciones para preparar los datos para su análisis. La columna **FECHA** se convierte a un dato tipo tiempo, permitiendo comparaciones temporales. Luego, **MATERIAL** se mapea con un diccionario, reemplazando códigos por descripciones (por ejemplo, “B” se convierte en “Tubo metálico” y la letra “T” pasa a “Poste en Concreto”) mejorando la interpretabilidad de datos. A través de un unión, se integran datos adicionales de una tabla llamada “Parámetros_varios.xlsx”, aumentando la información disponible con columnas como **TIPO**, **CLASE**, **MATERIAL**, etc. Para algunos registros con valores específicos de **ASSEMBLY** como “FICTICIO”, “PORTICO” o “FACHADA”, se asignan longitudes predeterminadas en “LONG_APOYO”. Valores nulos en **TOWNER** se rellenan con la letra “C” y **VIEN_SEC** se transforma a tipo numérico, mientras que **VIEN_PRIM** usa un diccionario, para estandarizar los datos y facilitando análisis cuantitativo. Registros sin **TIPO** asignan valores de **ASSEMBLY** basados en el último registro de **CODE** con información, un paso crucial para corregir una alteración en el orden de los códigos en el último semestre de datos debido a errores humanos, restaurando así la integridad de la información.

La tercera tabla, **REDMT**, contiene información esencial sobre la red de media tensión y es transformada para su respectivo análisis. Primero, se seleccionan registros con valores específicos en **KVNOM**, centrando el análisis en niveles de tensión menores a 33KV. La columna **PHASES** se transforma utilizando un diccionario que reemplaza cada código numérico con la fase correspondiente, estandarizando así la información. con una unión con la tabla “Parámetros_varios.xlsx” se incorpora detalles adicionales sobre conductores y neutral, agregando información como **MATERIALCONDUCTOR**, **TIPOCONDUCTOR**, y **CALIBRECONDUCTOR**. Asimismo, la columna **OWNER** se limpia y estandariza: los nombres se normalizan, y se asigna el valor “C” a los registros relacionados con CHEC y otros operadores específicos, unificando el formato de propietarios en toda la tabla. Las fechas se convierten en datos tipo tiempo, y valores nulos se reemplazan con una fecha predeterminada. Se calcula **FECHA_ACT** como el máximo entre fechas, simplificando el análisis temporal. Información sobre capacidad y resistencia de conductores se añade para completar la tabla y aumentar la información disponible.

La cuarta tabla, **SWITCHES**, almacena datos sobre interruptores en la red de media tensión. Primero, se realiza un filtro en **KV** conservando solo registros con niveles específicos propios de red de media tensión, lo cual asegura que solo se incluyan interruptores relevantes para el análisis. **PHASES** se mapea por medio de un diccionario para la uniformidad de los datos. En **ASSEMBLY**, se reemplaza códigos con descripciones detalladas, como “Cortacircuito monofásico” o “Seccionador trifásico de operación con carga”, facilitando la comprensión. **STATE** se mapea y asigna etiquetas descriptivas, asignando etiquetas como “Normalmente abierto” o “Normalmente cerrado” a cada valor numérico, estandarizando el estado de cada interruptor. a continuación, Se convierte la columna **FECHA** a tipo tiempo, lo que permite un análisis temporal efectivo. Para registros donde **ASSEMBLY** está vacío, se asigna un valor basado en el último registro del mismo **CODE** con **ASSEMBLY** no vacío, lo que ayuda a completar datos faltantes y corrige inconsistencias.

Finalmente, la tabla **TRAFOS** se transforma de la siguiente manera. **PHASES** se mapea para garantizar uniformidad e interpretabilidad, mientras que para **OWNER1** se asigna las letras “C” o “P” según sea el caso. **TIPO_SUB** se actualiza proporcionando descripciones detalladas de cada tipo de subestación, como “Subestación de patio de alta y extra alta tensión” o “Subestaciones tipo pedestal”. Las fechas en **DATE_FAB**, **DATE_INST**, y **DATE_REM** se convierten a datos tipo tiempo, relleno de valores nulos con una fecha predeterminada. La columna **FECHA_ACT** se calcula como el máximo entre **DATE_REM** y **DATE_INST**, lo que permite identificar la fecha de última actividad de cada transformador. Luego, se eliminan estas columnas para optimizar el tamaño del archivo. Un unión con “Parámetros_varios.xlsx” aumenta la información disponible para esta tabla, por ejemplo, KVA y KV, excluyendo registros no pertinentes.

Para integrar las tablas **EVENTOS**, **TRAFOS**, y **SWITCHES**, se realiza un cruce que permite relacionar eventos específicos tanto con transformadores como con interruptores, aumentando la información disponible para el análisis de criticidad en la red de media tensión de CHEC. La tabla **EVENTOS** se carga, y las columnas **inicio** y **fin** se convierten a tipo tiempo, generando además una columna **FECHA_C** que convierte las fechas

de inicio en periodos mensuales. Esta agrupación temporal facilita el cruce con las fechas de fabricación y operación en las tablas **TRAFOS** y **SWITCHES**. En estas últimas tablas, la columna **FECHA** también se convierte a tipo tiempo, y se crea la columna **FECHA_C** en periodos mensuales para asegurar consistencia en el cruce. La columna **CODE** se renombra a **equipo_ope** en ambos casos, asegurando que cada equipo y mes sean únicos. Para el cruce, se filtran primero en **EVENTOS** los registros donde **tipo_equi_ope** es “transformador” y se realiza una unión con **TRAFOS** usando las columnas **equipo_ope** y **FECHA_C**, lo cual permite relacionar cada evento con su transformador correspondiente. Posteriormente, se filtran los eventos donde **tipo_equi_ope** es “interruptor” y se hace un cruce similar con la tabla **SWITCHES**, vinculando los eventos con los interruptores. En ambos casos, se conservan los índices originales de **EVENTOS** y se actualizan los valores con el resultado de los cruces, incorporando así la información de transformadores e interruptores en cada evento relevante. Este cruce múltiple da como resultado una tabla de **EVENTOS** con datos críticos tanto de transformadores como de interruptores.

4.2. Desarrollo del sistema de análisis estadístico

El sistema desarrollado es una herramienta interactiva destinada al análisis estadístico de eventos en equipos eléctricos, enfocándose en transformadores y switches. A continuación, se describen los pasos llevados a cabo en su desarrollo, así como los componentes principales y su funcionamiento.

4.2.1. Pasos en el desarrollo del sistema

1. Definición de objetivos y requisitos Se identificaron los objetivos del sistema, que incluyen la capacidad de analizar eventos en equipos eléctricos y proporcionar visualizaciones interactivas. Se recogieron requisitos específicos para asegurar que el sistema cumpliera con las necesidades de los usuarios.

2. Estructura de datos Se establecieron dos DataFrames principales para gestionar los registros de eventos:

- **Eventos trafos:** Este DataFrame contiene los registros de eventos relacionados con transformadores, permitiendo un análisis específico de su comportamiento y rendimiento.
- **Eventos switches:** Contiene los registros de eventos vinculados a switches, facilitando la evaluación de su operatividad y funcionamiento en diversas condiciones.

3. Desarrollo de la interfaz de usuario La interfaz del sistema fue diseñada para incluir cuatro elementos fundamentales:

1. **Selectores de criterios:** Se implementaron cuatro criterios fijos que permiten al usuario definir las condiciones del análisis.
2. **Selectores de sub-criterios:** Se ofrecieron opciones adicionales específicas para cada criterio principal, aumentando la flexibilidad en el filtrado de datos.
3. **Filtros dinámicos:** Se desarrollaron filtros que se adaptan al tipo de dato seleccionado, permitiendo una personalización efectiva del análisis.
4. **Visualizador de distribución:** Se incorporó un visualizador que utiliza el método de Estimación de Densidad Kernel (KDE) para mostrar la distribución de los datos seleccionados.

4. Implementación del sistema de filtrado Se implementó un sistema de filtrado multinivel que permite:

- **Filtrado categórico:** Para variables discretas, facilitando la selección de grupos específicos de datos.
- **Filtrado por rangos:** Para variables continuas, permitiendo al usuario establecer límites de interés.
- **Filtrado temporal:** Para variables de tipo fecha, asegurando que los análisis se centren en intervalos de tiempo relevantes.
- **Manejo de múltiples condiciones simultáneas:** Permitiendo la aplicación de varios filtros de manera conjunta, aumentando la especificidad del análisis.

5. Análisis de estimaciones condicionales El sistema implementa el concepto de probabilidad condicional, expresado matemáticamente como:

$$P(Y|X_1, X_2, \dots, X_n) \quad (1)$$

Donde:

- Y : es la variable objetivo de interés.
- X_1, X_2, \dots, X_n : son las condiciones de filtrado seleccionadas por el usuario.

4.3. Diseño y desarrollo del mockup

El diseño del mockup para el dashboard de criticidad de CHEC se llevó a cabo mediante un proceso estructurado que priorizó la identidad corporativa y la funcionalidad del sistema. Este proceso se dividió en varias etapas clave, asegurando una interfaz intuitiva y profesional que integra la visualización de datos críticos de la red eléctrica y un sistema de chat inteligente.

4.3.1. Pasos en el desarrollo del mockup

1. Definición de objetivos y requisitos Se establecieron los objetivos principales del dashboard, que incluían la visualización efectiva de datos críticos de la red eléctrica y la implementación de un sistema de chat inteligente. Se identificaron los requisitos específicos basados en las necesidades de los usuarios y las funcionalidades deseadas.

2. Diseño de la interfaz Se diseñó una interfaz centrada en la identidad corporativa de CHEC, utilizando su logotipo y una paleta de colores en tonos verdes. La interfaz fue estructurada para facilitar la navegación y la accesibilidad, con una separación clara de las funcionalidades.

3. Estructuración de pestañas El dashboard se organizó en dos pestañas principales:

- **Visualización de Mapas:**
 - **Sección superior:** Se incorporó un selector de rango de fechas para filtrar datos temporalmente.
 - **Sección inferior izquierda:** Se diseñó un mapa para la ubicación de equipos y elementos de la red, como switches y transformadores.
 - **Sección inferior derecha:** Se implementó una visualización de eventos críticos, como interrupciones y descargas eléctricas, para facilitar el análisis de incidencias.
- **Chat Inteligente:**
 - **Panel izquierdo - historial:** Se desarrolló un registro histórico de conversaciones para un acceso rápido a interacciones previas.
 - **Panel derecho - chat activo:** Se creó una interfaz de conversación activa, con cuadros de diálogo diferenciados por avatares para el usuario y el chatbot.

4. Implementación y validación El desarrollo del mockup siguió una metodología ágil, permitiendo iteraciones constantes y retroalimentación. Esto garantizó que tanto las funcionalidades como el diseño visual estuvieran alineados con los estándares de la empresa.

5. Proyección de tecnologías Se evaluaron y seleccionaron tecnologías orientadas a la creación de interfaces web interactivas y sistemas de comunicación inteligente. Esto incluyó frameworks de visualización de datos geográficos y bibliotecas de deep learning para el sistema de chat, asegurando una solución robusta y adaptable a las necesidades operativas de CHEC.

4.4. Implementación de TabNet como Autoencoder

TabNet es un modelo de atención secuencial diseñado para procesar datos tabulares de manera eficiente y explicar la relevancia de cada característica en la predicción. En el contexto de un autoencoder, TabNet puede aplicarse para aprender una representación comprimida de las características de entrada, seleccionando solo las más relevantes en cada paso secuencial.

4.4.1. Estructura Matemática de TabNet

TabNet emplea una serie de pasos de decisión, en cada uno de los cuales se seleccionan características usando máscaras de atención generadas a partir de las salidas de atención y la importancia acumulada de pasos anteriores.

1. **Máscara de Selección de Características** $M[i]$: En cada paso i , TabNet genera una máscara de atención, $M[i]$, que se calcula como:

$$M[i] = \text{sparsemax}(P[i-1] \cdot h_i(a[i-1]))$$

donde:

- $P[i-1]$ es la importancia acumulada de cada característica hasta el paso $i-1$, actuando como una memoria histórica.
- $h_i(a[i-1])$ es una función no lineal de la salida del paso anterior, adaptando la relevancia de las características seleccionadas.
- **Sparsemax** asegura que solo algunas posiciones tengan valores distintos de cero, permitiendo que el modelo enfoque su atención en un subconjunto reducido de características.

2. **Escala de Importancia Acumulada** $P[i]$: La importancia acumulada se actualiza en cada paso según:

$$P[i] = P[i-1] \cdot (\gamma - M[i])$$

donde:

- γ : Es un hiperparámetro que controla la frecuencia con la que una característica puede seleccionarse en múltiples pasos.
- $M[i]$: Es la máscara generada en el paso i , que reduce la importancia de las características seleccionadas, incentivando la exploración de otras en pasos sucesivos.

3. **Sparsemax**: La función Sparsemax convierte la máscara preliminar en una máscara dispersa. La ecuación es:

$$\text{sparsemax}(z) = \max(z - \tau(z), 0)$$

donde $\tau(z)$ se ajusta para que la suma de las posiciones activadas sea 1, permitiendo una selección limitada de características.

4.4.2. Estructura del Autoencoder

En la configuración de autoencoder, TabNet se adapta para realizar dos tareas: **codificar** (encoder) las características de entrada en una representación comprimida y **decodificar** (decoder) esta representación para reconstruir la entrada original. Esta implementación consiste en tres módulos principales: **TabNetEncoder**, **TabNetDecoder**, y **TabNetAutoencoder**.

TabNetEncoder El encoder está diseñado para seleccionar dinámicamente las características relevantes en cada paso de decisión, permitiendo aprender una representación comprimida que retenga solo la información esencial.

Parámetros del Encoder:

- **units**: Define la dimensión de la salida codificada en cada paso.

- **n_steps**: Controla el número de pasos de decisión secuenciales.
- **n_features**: Dimensión de las características procesadas en cada transformación.
- **gamma**: Controla la frecuencia de reutilización de características en múltiples pasos.
- **sparsity**: Término de regularización que incentiva una máscara dispersa.

Componentes Principales:

- **Batch Normalization**: Estandariza las características de entrada.
- **Shared Feature Transformer**: Bloque transformador de características aplicado inicialmente a las entradas.
- **Pasos de Decisión (TabNetStep)**: Cada paso aplica transformaciones de características y genera una máscara de atención.

TabNetDecoder El decoder reconstruye la entrada original a partir de la representación comprimida generada por el encoder.

Parámetros del Decoder:

- **units**: Dimensión de la salida del decoder en cada paso.
- **n_steps**: Número de pasos de transformación en el decoder.
- **n_features**: Dimensión de las características procesadas en cada paso.

Componentes Principales:

- **Shared Feature Transformer**: Transformaciones iniciales sobre la representación comprimida.
- **FeatureTransformerBlocks**: Transformaciones adicionales en cada paso.
- **Capa Densa para Reconstrucción**: Capa que aproxima la reconstrucción final de la entrada.

TabNetAutoencoder Este módulo integra el encoder y el decoder para formar el autoencoder completo.

4.4.3. Función de Costo

El modelo utiliza la **Binary Crossentropy (BCE)** como función de costo principal, que mide la divergencia entre las distribuciones de la entrada y la salida reconstruida.

1. Binary Crossentropy (BCE):

$$\text{BCE} = -\frac{1}{N} \sum_{i=1}^N (y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i))$$

donde y_i representa la entrada y \hat{y}_i la salida reconstruida.

2. Regularización de Sparsity: Incentiva una selección dispersa de características en cada paso:

$$\text{Sparsity Loss} = \lambda \sum_{i=1}^{N_{\text{steps}}} \sum_{j=1}^d -M[i]_j \log(M[i]_j + \epsilon)$$

4.4.4. Entrenamiento y Optimización

El autoencoder se entrena utilizando **Adam** como optimizador con una tasa de apr

4.5. Desarrollo de Sistema de Recomendación y Consultas Mediante Técnicas de Lenguaje Natural

Se ha desarrollado un chatbot que permite a los usuarios de CHEC consultar información relevante acerca de los sistemas eléctricos de nivel de tensión 2, accediendo tanto a datos estructurados como a documentos normativos no estructurados. Este sistema facilita la obtención de respuestas rápidas y precisas sobre temas técnicos específicos, como las interrupciones en transformadores y switches, y el cumplimiento de normativas, como el RETIE (Reglamento Técnico de Instalaciones Eléctricas), integrando datos clave en un solo punto de consulta.

4.5.1. Arquitectura del Sistema

El chatbot desarrollado se basa en una arquitectura RAG (Retrieval-Augmented Generation), que permite combinar la generación de respuestas con la recuperación de información de bases de datos y documentos. Este enfoque es particularmente efectivo en situaciones donde es necesario acceder a datos específicos almacenados en tablas estructuradas o en documentos no estructurados, permitiendo obtener información precisa y contextualizada en cada consulta.

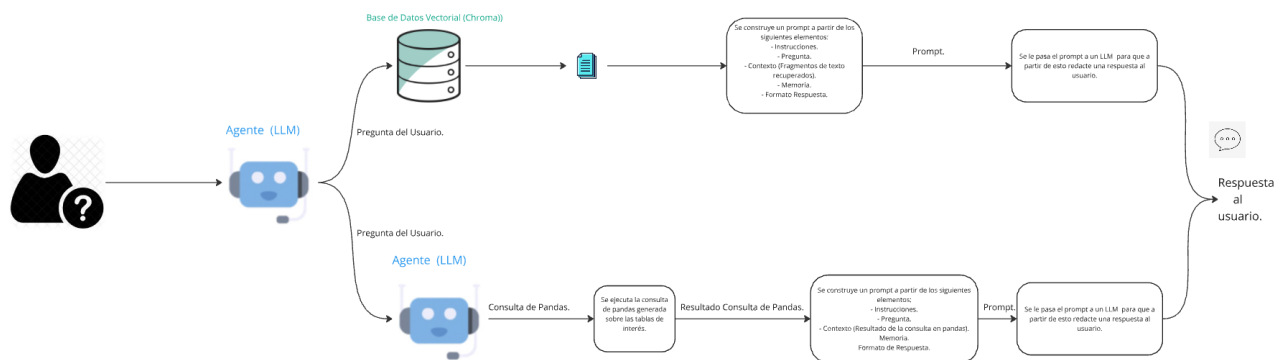


Figura 1: Arquitectura del Sistema de Recomendación y Consultas Mediante Técnicas de Lenguaje Natural

5. Resultados

5.1. Gráficas y visualización de distribuciones condicionales para eventos en equipos eléctricos

El sistema desarrollado se presenta como una herramienta interactiva diseñada para el análisis estadístico de eventos en equipos eléctricos, centrándose específicamente en transformadores y switches. Su funcionalidad principal es permitir la exploración y visualización de distribuciones probabilísticas condicionales, facilitando así el descubrimiento de patrones y relaciones relevantes entre diferentes variables de interés.

La estructura del sistema incluye dos DataFrames principales: uno para eventos relacionados con transformadores y otro para eventos relacionados con switches. Esto permite una organización clara y segmentada de los datos, asegurando que el análisis sea específico y relevante para cada tipo de equipo.

La interfaz de usuario está diseñada para ser intuitiva y funcional, incorporando elementos como selectores de criterios y sub-criterios, así como filtros dinámicos adaptados a los tipos de datos disponibles. La visualización de las distribuciones se realiza mediante la estimación de densidad kernel (KDE), lo que proporciona una representación clara y efectiva de la información.

5.1.1. Fundamentos del diseño y funcionalidades

La elección de la arquitectura y componentes del sistema se basa en varias consideraciones clave:

- **Estructura de datos clara:** La separación de eventos en diferentes DataFrames permite un manejo más eficiente de la información, facilitando el análisis de cada tipo de equipo sin confusiones ni mezclas de datos.
- **Interfaz intuitiva:** Los elementos de la interfaz, como los selectores de criterios y filtros, están diseñados para maximizar la usabilidad. Esto es fundamental para los usuarios que buscan resultados rápidos y precisos sin una curva de aprendizaje pronunciada.
- **Sistema de filtrado multinivel:** La implementación de un filtrado categórico, por rangos y temporal, permite al usuario adaptar el análisis a sus necesidades específicas, asegurando que se pueda trabajar con datos relevantes y significativos.
- **Uso de KDE:** La elección de la Estimación de Densidad Kernel es crítica, ya que ofrece ventajas en la suavización de distribuciones, manejo de datos dispersos y comparación entre diferentes condiciones de filtrado. Esto permite una interpretación más rica y significativa de las distribuciones observadas.

Se eligió utilizar el método KDE debido a sus múltiples ventajas:

1. **Suavizado de distribuciones:** Proporciona una estimación suavizada de la distribución de probabilidad, facilitando la identificación de patrones subyacentes.
2. **Manejo de datos dispersos:** Permite mantener una visualización significativa de la distribución en situaciones donde los datos son escasos.
3. **Comparabilidad:** Facilita la comparación de distribuciones bajo diferentes condiciones de filtrado, normalizando las densidades.

La visualización de la distribución se implementa utilizando la biblioteca de Python Seaborn, mediante el siguiente código:

```
1 sns.histplot(data, kde=True, bins=30)
```

- **Visualización clara:** Las características de la visualización incluyen:

- Uso de 30 bins para equilibrar el detalle y el suavizado en la representación gráfica.
- Superposición de la curva KDE a los histogramas para facilitar la interpretación de la densidad.

5.1.2. Resultados de la interfaz de usuario

La figura 2 muestra el panel de selección de filtros, que permite a los usuarios establecer criterios y sub-criterios para la exploración de los eventos en transformadores y switches. Este panel es fundamental para facilitar la personalización del análisis y garantizar que los usuarios puedan enfocarse en los datos más relevantes para sus necesidades.

Criterio 1:	Eventos-Trafos	Sub-Criterio 1:	inicio	Desde:	02/01/2019	Hasta:	17/04/2024
Criterio 2:	Eventos-Trafos	Sub-Criterio 2:	tipo_duracion	Valor:	> 3 min		
Criterio 3:	Eventos-Trafos	Sub-Criterio 3:	PHASES	Operador:	>	Valor:	1
Criterio 4:	Eventos-Trafos	Sub-Criterio 4:	cto_equi_ope	Valor:	VMA23L15		

Confirmar criterios

Variable objetivo: duracion_h

Confirmar variable

Figura 2: Panel de selección de filtros a partir del criterio, subcriterios y los filtros para cada tipo

La figura 3 ilustra cómo las distribuciones son presentadas al usuario, permitiendo una visualización clara y efectiva de las probabilidades condicionales. Esto es fundamental para la toma de decisiones informadas y la identificación de anomalías en los eventos registrados.

P(duracion_h|Eventos-Trafos , inicio = 2019-01-01 2024-04-17 , tipo_duracion = > 3 min , PHASES > 1.0 , cto_equi_ope = VMA23L15)

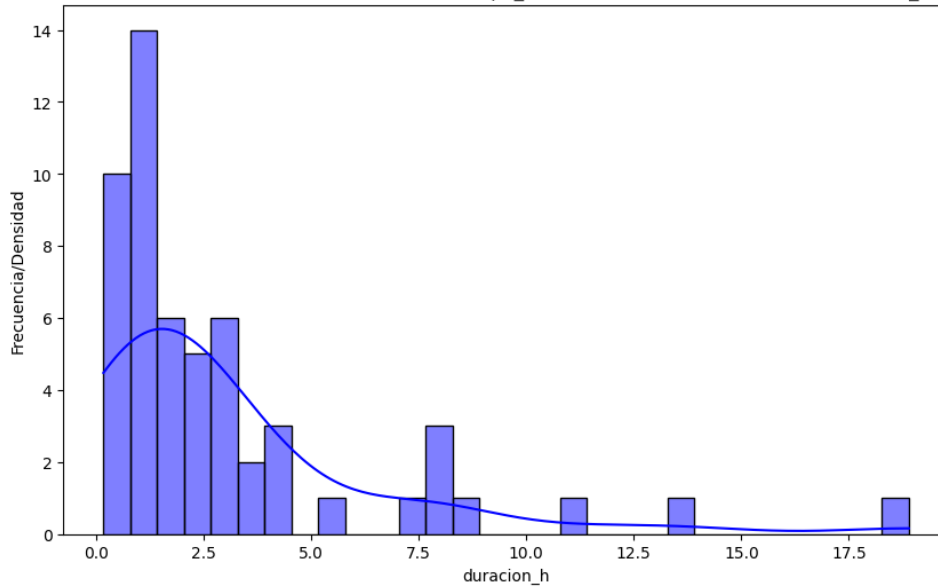


Figura 3: Visualización de la distribución y probabilidad condicional mediante histogramas y KDE

5.2. Mockup del dashboard

El mockup preliminar del Dashboard de Criticidad de CHEC se ha diseñado para facilitar la visualización de datos críticos de la red eléctrica y la integración de un sistema de chat inteligente. Este diseño se organiza en dos pestañas principales: *Visualización de Mapas* y *Chat Inteligente*, cada una diseñada con un enfoque claro en la funcionalidad y usabilidad.

5.2.1. Pestaña de visualización de mapas

La pestaña de *Visualización de mapas* se estructura en tres secciones principales:

- **Sección superior:** Contiene un selector de rango de fechas que permite a los usuarios filtrar la información temporalmente, garantizando que los datos mostrados sean relevantes al período seleccionado.
- **Sección inferior izquierda:** Presenta un mapa interactivo que muestra la ubicación de los equipos y elementos de la red CHEC, como switches, transformadores y apoyos, permitiendo así la monitorización en tiempo real y la ubicación precisa de estos componentes.
- **Sección inferior derecha:** Visualiza eventos críticos, como interrupciones y descargas eléctricas, proporcionando una herramienta valiosa para el análisis de incidencias en la infraestructura, permitiendo una respuesta rápida ante situaciones adversas.

5.2.2. Pestaña de chat inteligente

La pestaña de *Chat Inteligente* está dividida en dos secciones:

- **Panel izquierdo - historial:** Muestra un registro histórico de conversaciones, identificando cada chat con nombres cortos y representativos para facilitar el acceso rápido a interacciones previas, lo que mejora la eficiencia del usuario en la consulta de información.
- **Panel derecho - chat activo:** Contiene la conversación en tiempo real, con una barra superior que identifica la conversación actual y cuadros de diálogo visualmente diferenciados. Los avatares del usuario CHEC y del chatbot facilitan la interacción, proporcionando una experiencia de usuario clara y dinámica.

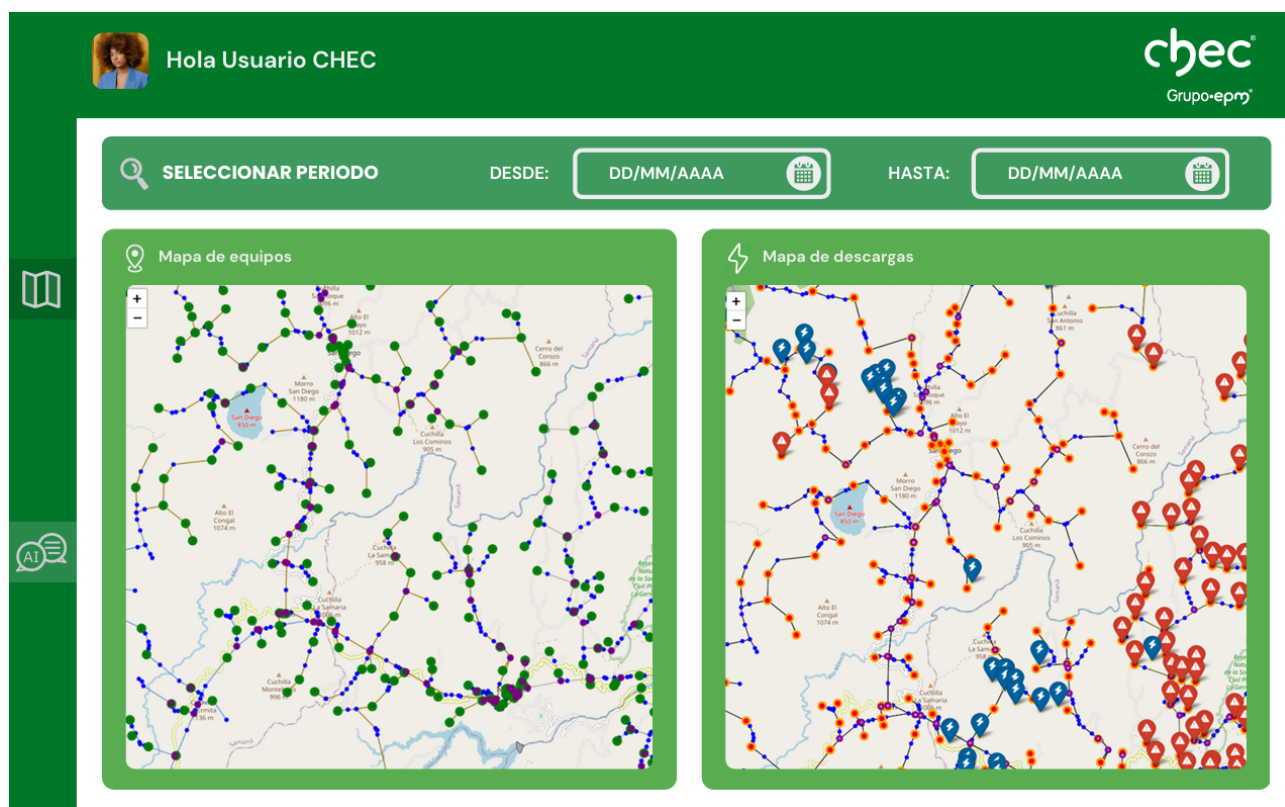


Figura 4: Pestaña con la visualización de mapas de equipos y elementos, y descargas eléctricas e interrupciones

5.2.3. Fundamentos del diseño

El diseño del dashboard se fundamenta con base a varios factores clave:

- **Identidad corporativa:** Se ha utilizado el logotipo oficial de CHEC y una paleta de colores basada en tonos verdes para asegurar la coherencia con la imagen institucional de la empresa. Esto no solo refuerza la identidad de marca, sino que también genera confianza en los usuarios.
- **Interfaz intuitiva:** La separación clara de funcionalidades y la consistencia visual en ambas pestañas permiten una navegación fluida y fácil, lo que es fundamental para un dashboard que maneja datos críticos. Esto ayuda a los usuarios a localizar rápidamente la información necesaria sin complicaciones.
- **Visualización efectiva de datos:** La estructura de las secciones permite que los datos críticos sean visualizados de manera eficiente. La integración de mapas y un sistema de chat inteligente en un solo dashboard proporciona un enfoque integral para la gestión de la red eléctrica, facilitando la toma de decisiones informadas y oportunas.
- **Facilidad de acceso a la información:** La disposición de los elementos y la inclusión de un historial de conversaciones en la pestaña de chat permiten a los usuarios acceder rápidamente a información relevante y previa, optimizando el flujo de trabajo y la interacción con el sistema.

5.3. Análisis de Resultados de TabNet como Autoencoder

La implementación del modelo **TabNet** como autoencoder en la base de datos de interrupciones de redes eléctricas, enfocada en eventos donde el elemento afectado fue un interruptor, permitió un análisis detallado de la criticidad y la identificación de patrones relevantes en los datos. A continuación, se describen los resultados obtenidos y su interpretación en el contexto del análisis de criticidad de la red.

Capacidad de Reconstrucción y Precisión del Modelo

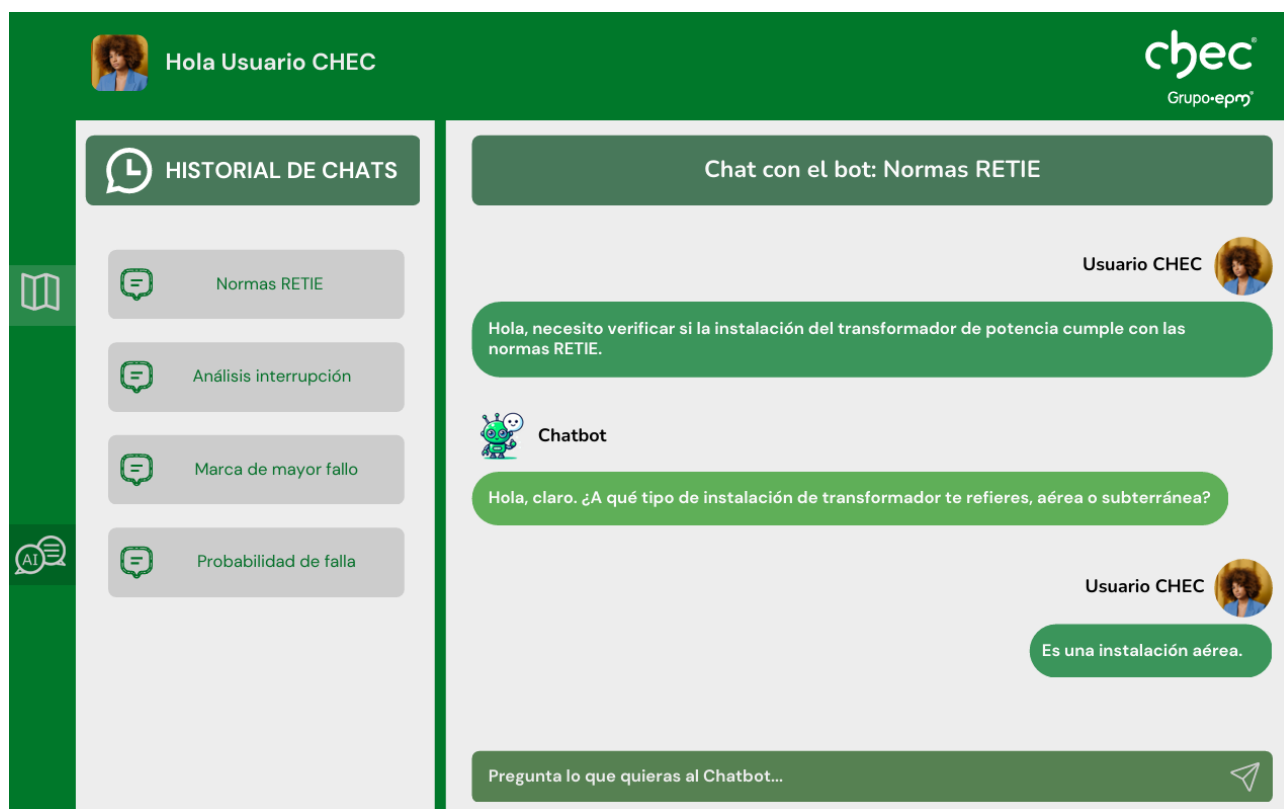


Figura 5: Pestaña con el historial de conversaciones y el chat con el bot

La Figura 6 muestra la pérdida (error) durante el entrenamiento del autoencoder TabNet a lo largo de las épocas.

La gráfica revela una rápida convergencia de la pérdida hacia valores cercanos a cero, estabilizándose después de aproximadamente 20 épocas. Esto indica que el autoencoder logra una reconstrucción precisa de las características relevantes de las interrupciones, manteniendo una mínima pérdida de información. Este comportamiento destaca la eficiencia del modelo en captar patrones complejos de los eventos de interrupción, lo cual es fundamental para el análisis de criticidad. La precisión en la reconstrucción respalda la validez de los datos transformados, proporcionando una base sólida para identificar componentes críticos en la red.

Identificación de Atributos Clave para la Criticidad

La Figura 7 presenta la importancia promedio de cada característica según la selección realizada por el autoencoder TabNet.

La gráfica de importancia destaca ciertos atributos, tales como `tipo_duracion`, `CNT_TRAFOS_AFEC`, `ASSEMBLY`, `FPARENT` y `XPOS`, como características con un peso significativo en la modelación de los eventos de interrupción. Estos atributos son esenciales para explicar la variabilidad en los datos y para capturar patrones relevantes en la criticidad de los eventos. La identificación de estos atributos clave permite enfocar el análisis en los factores determinantes para la robustez de la red, facilitando la priorización en la gestión de activos críticos y el mantenimiento. Por otro lado, los atributos con baja relevancia sugieren la posibilidad de optimizar los datos de entrada y reducir la dimensionalidad sin comprometer la precisión del diagnóstico.

Adaptabilidad del Modelo en la Selección de Características por Contexto de Interrupción

La Figura 8 muestra un mapa de calor que representa la importancia de cada característica para cada muestra específica en el conjunto de datos.

El mapa de calor refleja una variabilidad significativa en la importancia de cada atributo dependiendo de la muestra, lo que indica que el modelo adapta dinámicamente la selección de características en función del contexto específico de cada interrupción. Características como `CNT_TRAFOS_AFEC` y `FPARENT` muestran variabilidad en su importancia, lo cual subraya la capacidad del autoencoder TabNet para ajustar su atención a las condiciones particulares de cada evento. Esta adaptabilidad es crucial para el análisis de criticidad, ya que permite una

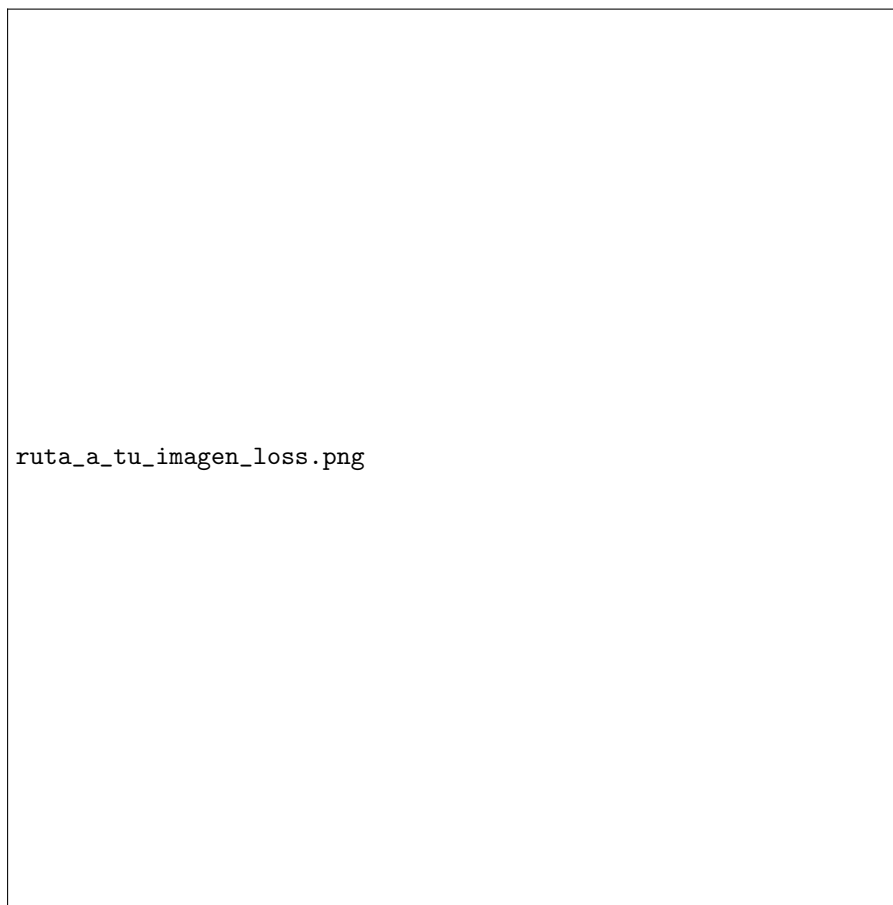


Figura 6: Pérdida durante el entrenamiento del modelo.

evaluación diferenciada de los factores de riesgo específicos para cada situación, mejorando la precisión en la interpretación de los eventos críticos y potenciando el diagnóstico bajo condiciones variables de la red eléctrica.

Aplicaciones para el Diagnóstico de Criticidad y Gestión de Activos

La implementación de TabNet como autoencoder proporciona una herramienta robusta para el análisis de criticidad en redes eléctricas. La capacidad del modelo para reconstruir los datos con precisión y para seleccionar características relevantes en función de cada muestra respalda su utilidad en la identificación de elementos críticos que afectan la estabilidad de la red.

La información obtenida a partir de este análisis puede aplicarse directamente en estrategias de mantenimiento preventivo y predictivo, permitiendo una detección temprana de componentes con alto nivel de riesgo. Además, facilita la priorización en la gestión de activos y la planificación de intervenciones focalizadas en los puntos críticos de la infraestructura de la red.

En conclusión, los resultados obtenidos demuestran que el modelo TabNet aplicado como autoencoder cumple con los objetivos propuestos para el diagnóstico de criticidad en redes eléctricas. Este avance contribuye a mejorar la confiabilidad de la red, permitiendo a la empresa tomar decisiones basadas en datos para optimizar su estrategia de mantenimiento y gestión de activos.

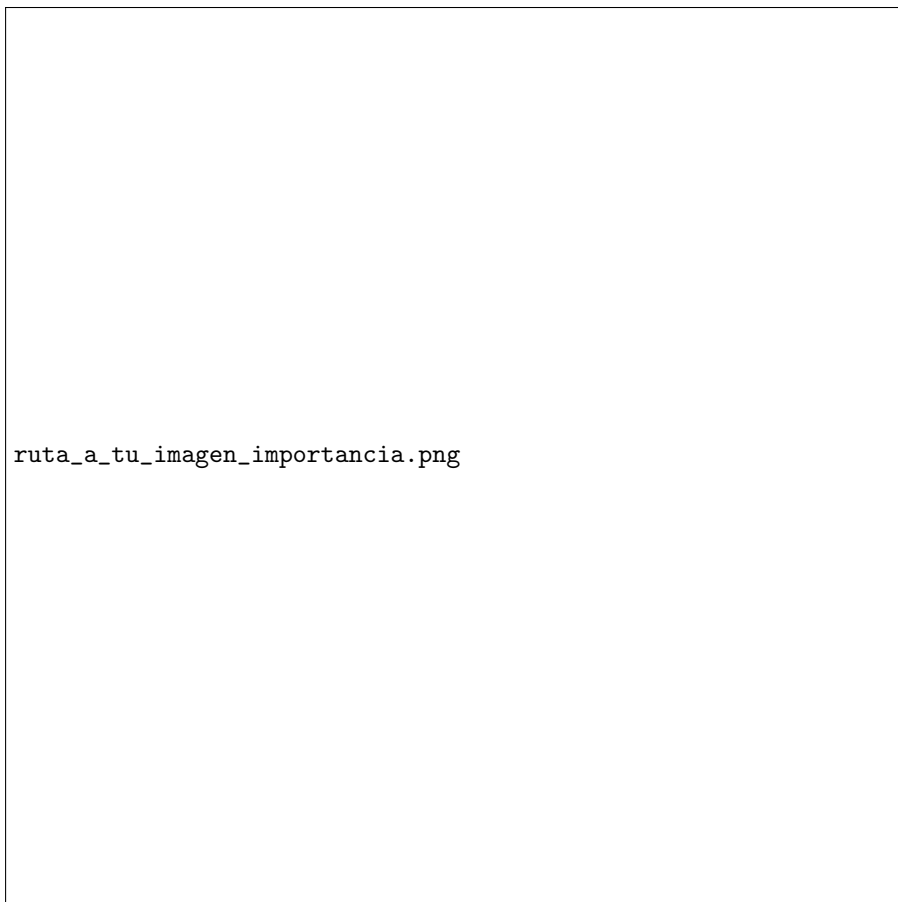


Figura 7: Importancia promedio de las características en el modelo.



Figura 8: Mapa de calor de la importancia de las características por muestra.