



UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE CUENCA
CARRERA DE INGENIERÍA ELÉCTRICA

PREDICCIÓN DE INTERRUPCIONES DE SERVICIO EN EL SISTEMA DE
DISTRIBUCIÓN DE LA EMPRESA ELÉCTRICA REGIONAL CENTRO SUR C.A.
UTILIZANDO MACHINE LEARNING

Trabajo de titulación previo a la obtención
del título de Ingeniera Eléctrica

AUTOR: VIANCA DANIELA ESPINOZA PESANTEZ

TUTOR: ING. EDGAR ANTONIO BARRAGÁN ESCANDÓN, Ph.D.

Cuenca - Ecuador

2022

**CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE
TITULACIÓN**

Yo, Vianca Daniela Espinoza Pesántez con documento de identificación N° 0106420904 manifiesto que:

Soy la autora y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Cuenca, 22 de abril del 2022

Atentamente,



Vianca Daniela Espinoza Pesántez

0106420904

**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE
TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Vianca Daniela Espinoza Pesántez con documento de identificación N° 0106420904, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autora del Proyecto Técnico: “Predicción de interrupciones de servicio en el sistema de distribución de la Empresa Eléctrica Regional Centro Sur C.A. utilizando machine learning” el cual ha sido desarrollado para optar por el título de: Ingeniera Eléctrica en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Cuenca, 22 de abril del 2022

Atentamente,



Vianca Daniela Espinoza Pesántez
0106420904

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Edgar Antonio Barragán Escandón con documento de identificación N° 0102516457, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: PREDICCIÓN DE INTERRUPCIONES DE SERVICIO EN EL SISTEMA DE DISTRIBUCIÓN DE LA EMPRESA ELÉCTRICA REGIONAL CENTRO SUR C.A. UTILIZANDO MACHINE LEARNING, realizado por Vianca Daniela Espinoza Pesántez con documento de identificación N° 0106420904, obteniendo como resultado final el trabajo de titulación bajo la opción Proyecto Técnico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Cuenca, 22 de abril del 2022

Atentamente,



Ing. Edgar Antonio Barragán Escandón

0102516457

AGRADECIMIENTOS

Agradezco a Dios por la vida, a toda mi familia, y amigos
Quienes han estado presente en este proceso y me han
Brindado el apoyo a lo largo de mi vida.

Un agradecimiento especial a Eduardo quien me ha brindado su apoyo
Y ayuda en este proceso y en mi vida.

A los ingenieros Edwin Lema, Sergio Zambrano, y Antonio Barragán,
tutores de este proyecto, un agradecimiento sincero por su apoyo,
guía y colaboración con la orientación y desarrollo del proyecto de Titulación.

A los profesores de la Universidad Politécnica Salesiana
por su eficiencia en la educación al transmitirnos no
solo conocimientos sino valores para la vida.

A la empresa Eléctrica Regional Centro Sur, por el apoyo brindado y el apoyo a la
investigación, que supieron abrir las puertas y apoyar en todo momento.

Vianca Espinoza.

DEDICATORIA

A mis papas Fabricio y Shajaira, que siempre han apoyado
Y me han brindado todo para que no me falte nada.

A mis abuelos Blanca, German, Vicente y Esilda que me han sabido inspirar a ser mejor
en todos los aspectos de mi vida y han estado ahí siempre que los necesito.

A todos mis tíos, de igual manera que han sido como hermanos
y quienes me han enseñado a ser mejor persona, a luchar por los sueños,
a ser perseverante y constante.

Y a mi querida hermana por ser mi lucha constante para enseñarle
Lo mejor, y por inspirarme a ser mejor persona,
por darme siempre su apoyo y compañía en los buenos y
malos momentos.

Vianca Espinoza.

**ESTE TRABAJO DE TITULACIÓN SE HA DESARROLLADO DENTRO DEL
CONVENIO ENTRE LA UNIVERSIDAD POLITECNICA SALESIANA Y LA
EMPRESA ELÉCTRICA REGIONAL CENTRO SUR.**

RESUMEN

Este trabajo describe el desarrollo y aplicación de modelos relacionados con la predicción de cortes no planificados en el sistema de distribución de la Empresa Eléctrica Regional CENTROSUR C.A. utilizando aprendizaje automático mediante el análisis y procesamiento de datos con el software Python™.

El modelo enfocado en la predicción de datos ocupará las herramientas disponibles de aprendizaje de máquina. Estas herramientas están menos exploradas en la industria de la distribución de energía; tanto a nivel nacional como internacional.

La variable extrínseca para la interrupción del servicio de la Empresa Eléctrica Centro Sur C.A. se obtuvo de estaciones meteorológicas de la ciudad de Cuenca. El aprendizaje necesita información histórica para predecir interrupciones. El resultado del modelo predictivo es comparado con la empresa distribuidora efectuando una interpolación por medio de, distancia inversa ponderada IDW (Inverse Distance Weighting).

El modelo de predicción se centrará únicamente en el área metropolitana de Cuenca, que incluye ocho subestaciones. Se emplea lenguaje de programación Python™ para establecer puntos de interpolación entre las subestaciones y un área delimitada, mediante el entrenamiento de una red neuronal que analiza la evolución de los coeficientes de proximidad espacial a lo largo del tiempo a nivel de cuadrícula geográfica.

Palabras clave: Redes neuronales artificiales, lógica difusa, machine learning, interrupciones, energía eléctrica.

ABSTRACT

This work describes the development and application of models related to the prediction of unplanned outages in the distribution system of Empresa Eléctrica Regional CENTROSUR C.A. using machine learning by analyzing and processing data with Python™ software.

The model focused on data prediction will use the available machine learning tools. These tools are less explored in the power distribution industry; both nationally and internationally.

The extrinsic variable for the interruption of the service of Empresa Eléctrica Centro Sur C.A. It was obtained from meteorological stations in the city of Cuenca. Learning needs historical information to predict interruptions. The result of the predictive model is compared with the distribution company, performing an interpolation by means of inverse distance weighted IDW (Inverse Distance Weighting).

The prediction model will only focus on the Cuenca metropolitan area, which includes eight substations. Python™ programming language is used to establish interpolation points between the substations and a delimited area, by training a neural network that analyzes the evolution of spatial proximity coefficients over time at the geographic grid level.

Keywords: Artificial neural networks, Fuzzy logic, Machine learning, Interruptions, Electric Power.

ÍNDICE

RESUMEN.....	5
ABSTRACT	6
I CAPÍTULO.....	13
GENERALIDADES	13
1.1 ESTADO DEL ARTE.....	13
1.2 INTRODUCCIÓN	14
1.3 ANTECEDENTES.....	15
1.4 JUSTIFICACIÓN.....	16
1.5 ALCANCE	16
1.6 OBJETIVOS	16
1.6.1 OBJETIVO GENERAL	16
1.6.2 OBJETIVO ESPECÍFICO	16
1.7 DESARROLLO DEL DOCUMENTO	17
II CAPÍTULO	18
INTERRUPCIONES EN LOS SISTEMAS DE DISTRIBUCIÓN: GENERALIDADES Y CONCEPTOS BÁSICOS.....	18
2.1 SISTEMA DE DISTRIBUCIÓN ELÉCTRICA	18
2.2 INTERRUPCIONES EN EL SISTEMA DE DISTRIBUCIÓN ELÉCTRICA	19
2.3 INTERRUPCIONES.....	20
2.3.1 Interrupción preventiva	20
2.3.2 Interrupción forzada	20
2.3.3 Interrupción programada	20
2.3.4 Interrupción no programada	20
2.4 SISTEMA ELÉCTRICO CENTROSUR	24
2.4.1 ANÁLISIS DE DATOS DE LA CENTROSUR.....	24
2.5 NORMAS DE CALIDAD EN EL SERVICIO ELÉCTRICO DE DISTRIBUCIÓN.....	25
2.6 DISPOSICIONES DE REGULACIONES VIGENTES.....	26
III CAPÍTULO	27
APRENDIZAJE PROFUNDO DE MÁQUINA Y HERRAMIENTAS DIGITALES.....	27
3.1 INTRODUCCIÓN	27
3.2 REDES NEURONALES.....	¡Error! Marcador no definido.
3.2.1 FUNDAMENTOS.....	¡Error! Marcador no definido.
3.2.2 ESTRUCTURA DE REDES NEURONALES.	31
3.2.3 CARACTERÍSTICAS DE LAS REDES NEURONALES.	32
3.2.4 CLASIFICACIÓN DE LAS REDES NEURONALES	32
3.2.5 ARQUITECTURA.....	33

3.2.6	CLASIFICACIÓN DE LOS MÉTODOS DE INTELIGENCIA ARTIFICIAL .	39
3.2.7	USO DE LA INTELIGENCIA ARTIFICIAL	39
3.3	HERRAMIENTAS DIGITALES.....	40
3.4	CRITERIOS PARA LA EVALUACIÓN DE LA METODOLOGÍA.....	41
3.4.1	PORCENTAJE DE ERROR MEDIO ABSOLUTO.....	41
3.4.2	ERROR MEDIO CUADRÁTICO (MSE-MEAN SQUARE ERROR)	42
3.4.3	COEFICIENTE KAPPA.....	42
3.4.4	FACTOR DE CORRELACIÓN DE PEARSON.....	42
3.4.5	R2	43
3.4.6	Métrica Recall_Score	43
3.4.7	Métrica f1_Score.....	43
3.5	INTERPOLACIÓN DE PONDERACIÓN DE DISTANCIA INVERSA.....	44
3.6	MODELOS PARA LA PREDICCIÓN DE FALLAS Y DEFINICIÓN DE LOS MODELOS	44
3.6.1	MLP REGRESSOR.....	44
3.6.2	MLP CLASSIFIER	45
3.7	K-MEANS.....	45
IV	CAPÍTULO.....	46
	METODOLOGÍA Y APLICACIÓN DEL MODELO	46
4.1	INTRODUCCIÓN	46
4.2	RECOLECCIÓN DE DATOS	46
4.2.1	INFORMACIÓN DE INTERRUPCIONES CENTROSUR.....	46
4.2.2	INFORMACIÓN DE ESTACIONES METEOROLÓGICAS	46
4.2.1	Presión atmosférica.....	46
4.2.2	Precipitación.....	46
4.2.3	Humedad relativa.....	47
4.2.4	Radiación solar global.....	47
4.2.5	Temperatura ambiente.....	47
4.2.6	Dirección del viento.....	47
4.2.7	Velocidad del tiempo.....	47
4.3	ANÁLISIS DE DATOS DE INGRESO	47
4.3.1	Correlación entre variables con la duración de la falla.....	48
4.4	MAPAS GEOGRÁFICOS	49
4.4.1	UBICACIÓN DE LAS ESTACIONES METEOROLÓGICAS	49
4.4.2	DEFINICIÓN DE ÁREA.....	49
4.5	METODOLOGÍAS APLICADAS, APLICACIÓN DEL MODELO.....	50

4.5.1	ANÁLISIS DE DATOS Y LISTADO DE VARIABLES DE ENTRADA Y SALIDA.	52
4.5.2	PREPROCESAMIENTO DE DATOS	52
4.5.3	AJUSTE DE DATOS.....	52
4.5.4	INTERPOLACIÓN DE VARIABLES METEREOLÓGICAS	54
4.5.5	MODELOS PARA LA PREDICCIÓN DE FALLAS Y DEFINICIÓN DE LOS MODELOS	55
4.5.6	K-MEANS.....	55
4.6	ENTRENAMIENTO.....	56
4.6.1	FUNCIONES DE ACTIVACIÓN Y OPTIMIZACIÓN.....	56
4.7	RESULTADOS.....	58
4.7.1	MLP REGRESSOR.....	58
4.7.2	MLP CLASIFIER	59
V	CAPÍTULO	64
	DISCUSIÓN	64
5.1	DISCUSIÓN.	64
VI	CAPÍTULO.....	65
	CONCLUSIONES Y RECOMENDACIONES.....	65
6.1	CONCLUSIONES.	65
6.2	RECOMENDACIONES	66
	BIBLIOGRAFÍA.....	67
	GLOSARIO.....	72

ÍNDICE DE FIGURAS

Figura 1. Configuración de un sistema eléctrico de potencia. Fuente:[23].	18
Figura 3. Neurona biológica. Fuente: [39].	¡Error! Marcador no definido.
Figura 4. Ejemplo básico de una estructura neuronal. Fuente:[40].	28
Figura 5. Secciones de una red neuronal.	29
Figura 6. Pesos numéricos de una neurona artificial [22].	29
Figura 7. Grafica de función bipolar [22].	¡Error! Marcador no definido.
Figura 8. Grafica de tangente hiperbólica [22].	30
Figura 9. Grafica de función sigmoide [22].	¡Error! Marcador no definido.
Figura 10. Grafica función Relú [22].	31
Figura 11. Ejemplo de una red neuronal. Fuente:[39].	32
Figura 12. Camino de la red para el entrenamiento Feedwars y Backward. Fuente: [41].	34
Figura 13. Región de decisión de un perceptrón simple. Fuente: [22]	35
Figura 14. Región de decisión de la RN con arquitectura PMC. Fuente: [22].	37
Figura 15. RNR y despliegue. Fuente: [22].	37
Figura 16. Clasificación de métodos. Fuente: [40].	39
Figura 17. Entorno gráfico de ArcGis®.	41
Figura 18. Interpolación ponderada de distancia Inversa. Fuente: [57].	44
Figura 19. Factor de correlación Pearson.	48
Figura 20. Correlación gráfica de las variables externas con respecto a la duración.	48
Figura 21. Ubicación de las estaciones Meteorológicas.	49
Figura 22. Mapa de Definición de Áreas de CENTROSUR, Área de Concesión, Azuay y Cobertura Subestaciones.	50
Figura 23. Flujograma del algoritmo de la red neuronal.	52
Figura 24. Interrupciones de la CENTROSUR dentro del Área de Cobertura Subestaciones.	53
Figura 25. Obtención de distancias mediante las cuadrículas y las estaciones.	54
Figura 26. Ponderación de distancia inversa en la velocidad del viento.	55
Figura 27. Punto de Codo.	56
Figura 28. Resultado del entrenamiento con la función de activación Identity y optimizador Sgd.	57
Figura 29. Resultado del entrenamiento con la función de activación Relú y optimizador Adam.	58
Figura 30. Resultado del Entrenamiento con MLP Regressor.	58

Figura 31. Resultado del Test con MLP Regressor.....	59
Figura 32. Distribución del tamaño de muestras para criterio de K-Means y de expertos de la CENTROSUR.....	60
Figura 33. Resultado del Test con MLP Classifier con K-Means.....	62
Figura 34. Representación de las interrupciones en el año 2015.....	76
Figura 35. Representación de las interrupciones en el año 2016.....	79
Figura 36 Representación de las interrupciones en el año 2017.....	82
Figura 37. Representación de las interrupciones en el año 2018.....	85
Figura 38 Representación de las interrupciones en el año 2019.....	88
Figura 39. Mapa de calor de la Presión Atmosférica.....	91
Figura 40. Mapa de Calor de la Precipitación.....	92
Figura 41. Mapa de Calor de la Radiación Solar.....	93
Figura 42. Mapa de calor de la Humedad.....	94
Figura 43. Mapa de calor de la Temperatura.....	95
Figura 44. Mapa de calor de la Dirección del Viento.....	96
Figura 45. Mapa de calor de la Velocidad del Viento.....	97

INDICE DE TABLAS

Tabla 1. Clasificación de interrupciones. Fuente: [25].	20
Tabla 2. Clasificación de los tipos de cortocircuitos. Fuente: [29].	21
Tabla 3. Clasificación de los tipos de perturbaciones. Fuente:[31].	23
Tabla 4. Niveles de confianza para el valor de k.	42
Tabla 5. Entradas y salidas para diferentes casos de predicción.	52
Tabla 6. Posiciones de las horas para interpolar.	56
Tabla 7. Comparación de las funciones de activación y optimizadores.	57
Tabla 8. Resultados de los errores con la red MLP Regressor en el entrenamiento.	58
Tabla 9. Resultados de los errores con la red MLP Regressor en el test.	59
Tabla 10. División de grupos de interrupción de acuerdo a los criterios establecidos a la duración.	59
Tabla 11. Tamaño de muestra para los criterios establecidos.	60
Tabla 12. Resultados de los errores con la red MLP Classifier en el entrenamiento según el criterio de un experto.	60
Tabla 13. Resultados de los errores con la red MLP Classifier en el entrenamiento.	61
Tabla 14. Resultados de los errores con la red MLP Classifier en el entrenamiento según el criterio K-Means.	61
Tabla 15. Resultados de los errores con la red MLP Classifier en el test según el K-Means.	61
Tabla 16. Información de interrupción de Cortocircuito.	62

I CAPÍTULO

GENERALIDADES

1.1 ESTADO DEL ARTE

Los sistemas de distribución están constantemente bajo la amenaza de fallas de cortocircuito que causarían cortes de energía. Para mejorar la calidad operativa y la confiabilidad de los sistemas de distribución, los operadores del sistema deben lidiar con las interrupciones de manera oportuna. Por lo tanto, es de suma importancia localizar con precisión y despejar las fallas inmediatamente después de que ocurran, de modo que se pueda lograr una restauración rápida.

Hay dos desafíos principales para la localización de fallas en los sistemas de distribución con el mayor número de mediciones disponibles: primero, los métodos tradicionales de localización de fallas no pueden incorporar las mediciones de diferentes buses de manera flexible, especialmente cuando se toman en consideración las pérdidas de datos. En segundo lugar, para los enfoques tradicionales de aprendizaje automático, la topología de la red de distribución es difícil de modelar, y mucho menos la posibilidad de cambios en la topología [1].

Las técnicas de localización de fallas existentes en la literatura se pueden dividir en varias categorías, a saber, métodos basados en impedancia [1], métodos basados en huecos de voltaje [2], mapeo automático de cortes [3], basados en ondas viajeras métodos [4][5] y métodos basados en el aprendizaje automático [6].

Los modelos de aprendizaje automático se aprovechan para la localización de fallas en los sistemas de distribución. Utilizando las características espectrales de las mediciones posteriores a la falla, los datos con extracción de características se introducen en una red neuronal artificial (ANN) para la ubicación de la falla [7].

Se propone un modelo de regresión lineal para predecir el tiempo de duración de fallas en los sistemas de distribución, con el fin de superar la deficiencia y no homogeneidad de los datos de interrupción. La investigación muestra que el proceso para predecir las interrupciones puede mejorarse sustancialmente mediante el uso de datos meteorológicos para evaluar automáticamente el riesgo [8].

El autor en [9], presenta un modelo de regresión logística considerando los impactos de rayos en las líneas de transmisión utilizando técnicas de mitigación que se propusieron; para el análisis se consideraron los impactos de las variables externas de las estaciones meteorológicas, se tomaron en cuenta diferentes duraciones de fallas relacionadas con cortocircuito, incluidas fallas temporales y permanentes, obtenido como resultado un modelo de regresión logística clásico que genera la probabilidad de un resultado binario [9].

El autor en [10], determina el tiempo de duración basándose en un análisis de conjuntos extensos de datos históricos en tiempo real propone un estimador de pérdida de la tasa de fallas de las líneas aéreas, que compromete entre las estimaciones de la tasa de fallas individuales y agrupadas.

El autor en [11], propone una técnica que se enfoca en aprendizaje automático para predecir el tiempo de reparación y restauración, basándose en una combinación de modelos multiescala de bosque aleatorio y equilibrio bayesiano. Y la segunda parte utiliza Deep Neural Network (DNN) para predecir igual tiempos de reparación de los sistemas de distribución dañados, teniendo como resultado un rendimiento satisfactorio con un margen de error medio de 18 minutos en tiempo de reparación y de 2 horas en el tiempo de restauración.

En [12], se propone desarrollar dos modelos diferentes para los propósitos de predicción: usando un enfoque estadístico basado en algoritmos de series de tiempo para predecir las interrupciones

y empleando un enfoque basado en aprendizaje automático. Sin embargo, realizado varios análisis estadísticos (es decir, ajustando modelos lineales y analizando residuos y verificando los supuestos del modelo lineal) concluyó que los modelos lineales no son adecuados para estos problemas.

En [13], el enfoque propuesto incorpora información meteorológica, de activos y de interrupciones para caracterizar el riesgo en una sección de la red de distribución a través de la plataforma GIS, el estudio presentado permite la incorporación de parámetros climáticos, como la temperatura promedio y la incidencia de la iluminación, directamente asociados a voltajes térmicos y eléctricos.

El autor en [14], enfoca al uso innovador de los datos de correlación de análisis de big data para mejora de la gestión de cortes en las redes de distribución con datos meteorológicos, se examinó el marco de correlación del análisis de big data y describe varias fuentes de datos y su correlación con el modelo de red.

En el proyecto propuesto analizando varias alternativas de modelos se realiza un análisis que se utilizan datos de cortes de 6 años (interrupciones relacionadas con eventos climáticos); continuando con una segunda parte que se propone una técnica basada en el aprendizaje automático para predecir el tiempo después de las interrupciones [11]. Se entrena mediante el aprendizaje supervisado con datos de interrupciones de 5 años. Los datos del sexto año se utilizan para validar el rendimiento del modelo entrenado [11].

La principal contribución de este enfoque basado en datos es la propuesta de un modelo predictivo de aprendizaje automático para estimar los tiempos de duración en diferentes condiciones externas incorporando información meteorológica, y de interrupciones.

Se proporcionó un procedimiento detallado, que incluye varios pasos para tratar de manera efectiva la preparación de datos requerida, problemas prácticos asociados con el conjunto de datos de fallas del mundo real y la implementación de Multi-Layer Perceptron (MLP) [15].

Se presenta una mejora sobre un sistema de predicción de fallas, basado Logistic Regression y la aplicación de técnicas de predicción de eventos raros, capaz de trabajar con datasets dispersos y de alta dimensión [16]. Un problema muy común para este proyecto es la escasez de datos, la mayor parte de la matriz del conjunto de datos está compuesta por ceros (ausencias de eventos). Esta situación es complicada para la mayoría de los métodos de clasificación de aprendizaje automático habituales, donde tienden a sobre ajustarse o no estimar correctamente la variable objetivo [16].

1.2 INTRODUCCIÓN

La misión de la Empresa Eléctrica Regional Centro Sur C.A (CENTROSUR) es “Suministrar el servicio público de energía eléctrica para satisfacer las necesidades de la sociedad, cumpliendo estándares de calidad, con responsabilidad social, ambiental y económica, sobre la base de la gestión del talento humano e implantación de nuevas tecnologías” [17].

Los servicios de distribución de energía eléctrica deben proporcionar a los clientes energía confiable, segura y continua. El propósito es evitar confusiones al estimar la ocurrencia de un evento y determinar si el evento es a corto, mediano o largo plazo.

Para garantizar resultados confiables se utiliza como eje principal una de datos de las interrupciones que ocurren dentro de la empresa CENTROSUR en un periodo de tiempo. Estos eventos están asociados a variables externas y determinan si su ocurrencia se debe a factores que no están relacionados con el funcionamiento normal del sistema.

El nivel de exactitud en la predicción de la interrupción tiene como propósito garantizar que la empresa distribuidora efectúe una apropiada planificación de su sistema de distribución, sistemas de protección eléctrica o; centro de transformación. Estas condiciones de funcionamiento garantizan un servicio de alta calidad y aseguran que el cliente mantenga o mejore el nivel de confort.

La propuesta del proyecto es emplear Redes Neuronales, cuyos resultados serán valorados y comparados con las predicciones realizadas por CENTROSUR. Esta comparación permitirá validar la confiabilidad de los resultados.

Gracias al Sistema de Información Geográfica (GIS) de la Empresa Distribuidora, se puede implementar el algoritmo de tendencias dividiendo el área de concesión a nivel de cuadrículas homogéneas para luego llevar a un nivel cantonal, permitiendo la predicción de interrupciones de cada cantón.

Las predicciones de las interrupciones se desarrollan a nivel de cuadrículas uniformes que tendrán características asociadas de acuerdo con las variables externas obtenidas de las estaciones meteorológicas que cambian a lo largo de los años. Con las grillas se crearán mapas en base a estos factores que sirven para entrenar y evaluar el modelo con los elementos actuales, asignando una probabilidad de desarrollo para cada uno de esos factores.

1.3 ANTECEDENTES

Las empresas distribuidoras de energía eléctrica, están forzadas a perfeccionar la calidad del servicio eléctrico, los requisitos que requieren se disponen a través de normativas nacionales e internacionales como: IEC 6100-2-2/4 [18], y CENELEC 50160 [19], para garantizar la fiabilidad del sistema eléctrico se debe cumplir los requisitos de las normas técnicas y la satisfacción del cliente.

Estas empresas están tratando de encontrar soluciones prácticas que permitan predecir las interrupciones o administrar mejor sus tiempos de respuestas a las interrupciones [15]. Para identificar las principales causas y variables que conducen a las interrupciones en la red de distribución, se requiere una comprensión clara de los diferentes incidentes. En el caso particular la CENTROSUR se ha evidenciado un incremento en el número de interrupciones ocurridas en la red distribución en el área de concesión [20].

Se han registrado diferentes tipos de interrupciones, debido a diversas causas. Se ha establecido que algunos de estos trastornos están mediados por la presencia de factores específicos [15]. Por ejemplo: hora de la interrupción, lugar, causa. Así, en algunas zonas rurales y urbanas se producen averías periódicas por causas comunes como la vegetación, fenómenos climáticos, etc.[21].

En los últimos años, se han desarrollado herramientas muy efectivas para el aprendizaje profundo de máquina (APM), son una forma especial de algoritmo de aprendizaje de máquina (AAM). Estas herramientas de optimización se utilizan para resolver problemas específicos como predicciones de series de tiempo, procesamiento de imágenes, reconocimiento de voz. Dado que la mayoría de las computadoras actuales cuentan con un alto poder de cómputo y son muy eficientes desde el punto de vista computacional, el APM puede utilizar la potencia informática que se encuentra en el hardware moderno sin que el usuario tenga que utilizar recursos costosos y difíciles de encontrar. Además, esta herramienta también se utiliza para resolver problemas relacionados a sistemas eléctricos [22].

1.4 JUSTIFICACIÓN

El sistema de distribución de energía eléctrica es la parte básica dentro del sistema de distribución, suministra energía a diferentes usuarios. Además, las fallas en el sistema de distribución pueden causar suministros, transformadores, fallos de la red, e interrupciones en la entrega de servicios de energía en áreas urbanas y rurales [22]. Para garantizar la satisfacción del usuario, el suministro de energía no solo debe garantizarse en forma de interrupciones, sino que también debe cumplir con los estándares regulatorios locales e internacionales. La interferencia puede ser causada por cambios en la temperatura y la presión, radiación ultravioleta y ozono, viento, lluvia, humedad, smog o contaminación. [9]. Además, el sistema puede experimentar interrupciones y apagones en el servicio de energía que pueden generar graves pérdidas económicas.

Debido a la importancia de la resolución de problemas, las empresas de distribución de energía pueden realizar fácilmente encuestas para predecirlos. Esto tendrá como objetivo tomar decisiones informadas para implementar medidas de monitoreo y minimizar las posibles causas [22]. De esta forma, es posible localizar dónde se ha producido el daño e identificar los factores que más influyen en la reparación y recuperación energética de los elementos dañados [15] [11]. El objetivo de este estudio es generar modelos predictivos de fallas en las redes de distribución de energía a partir de otras fuentes, como los sistemas de informes meteorológicos, lo que ayudará a generar nuevos conocimientos en el campo [14].

1.5 ALCANCE

El trabajo actual es analizar y recopilar el historial anual de interrupciones de energía y factores externos en el área de concesión de CENTROSUR.

Los datos se utilizarán con la finalidad para predecir el tiempo de corte del servicio eléctrico en el sistema de distribución a lo largo el tiempo a nivel estatal utilizando un enfoque de Redes Neuronales Artificiales. También se analizarán las interrupciones que han ocurrido durante un periodo de 6 años para determinar la probabilidad de una nueva falla para cada cuadrícula que utilice la red neuronal.

Se utilizará herramientas computacionales como Python™ y ArcGIS® los cuales cuenta con bloques computacionales para la implementar algoritmos y asegurar el correcto manejo de las variables de entrada y salida.

El área de concesión de CENTROSUR se segmentará a nivel de cuadrícula geográficas uniforme a nivel cantonal, para permitir que los métodos analíticos clasifiquen las interrupciones según su causa y efecto a la que están sujetas.

Para la implementación del algoritmo se utilizarán los siguientes procedimientos: recolección y preprocesamiento de datos históricos, modelado, entrenamiento y prueba del algoritmo, presentación de los resultados obtenidos y mapas de calor, a evaluar.

1.6 OBJETIVOS

1.6.1 OBJETIVO GENERAL

Desarrollar un modelo para la predicción de fallas en el sistema de distribución de la Empresa Eléctrica Regional Centro Sur C.A. utilizando machine Learning.

1.6.2 OBJETIVO ESPECÍFICO

- Analizar el estado del arte mediante revisiones bibliográficas para la determinación de los conceptos importantes para este estudio.

- Preparar los datos históricos de las interrupciones del servicio y factores externos en el sistema eléctrico de distribución de la CENTROSUR.
- Construir y entrenar un modelo de predicción de fallas usando Python™ y librerías de Machine Learning.
- Evaluar el desempeño del modelo de predicción de fallas.

1.7 DESARROLLO DEL DOCUMENTO

El documento consta de 6 capítulos divididos de la siguiente manera:

En el capítulo 1, introducirá el enfoque global del contenido del proyecto, destacando sus generalidades a través de los siguientes puntos: Estado del arte, introducción, antecedentes, justificación, alcance, objetivos generales y específicos, estructura del documento.

El capítulo 2 abarcará el desarrollo del marco teórico, conceptos básicos, visión de las interrupciones, y descripción del sistema eléctrico de la empresa.

El capítulo 3 proporcionará la información sobre los diferentes métodos de predicción, centrándose en los métodos de inteligencia artificial utilizados, y las herramientas digitales que serán usadas para implementar algoritmos tales como ArcGis®, Machine Learning, y Python™.

El capítulo 4 describe el entrenamiento desarrollado en mapas de calor, clasificado según las variables, presentan fuentes e información de los datos obtenidos, y los factores utilizados para la creación de los mapas de predicción usando la herramienta GIS, se detallarán las metodologías implementadas a través de las etapas de: recolección de datos históricos, reprocesamiento de los datos, creación del diseño/modelo, entrenamiento, pruebas del algoritmo y representación de resultados obtenidos.

En el capítulo 5 cubre la discusión sobre el tema planteado.

Finalmente, se presentará una propuesta en el capítulo 6, con recomendaciones y conclusiones, sobre la metodología a utilizar, con base en la experiencia adquirida durante la implementación del proyecto.

II CAPÍTULO

INTERRUPCIONES EN LOS SISTEMAS DE DISTRIBUCIÓN: GENERALIDADES Y CONCEPTOS BÁSICOS

Este capítulo presenta lo que refiere la empresa eléctrica CENTROSUR. Se proporcionará información detallada sobre las fallas eléctricas basadas en IEEE 1782 que ocurran en los años 2014-2020, así como la ubicación geográfica y el momento de las fallas que ocurran durante el proceso. A la información de variables externas también se le asignarán brechas más precisas como las variables más importantes para las estaciones meteorológicas existentes dentro del área de operación de CENTROSUR o sus áreas más concentradas.

2.1 SISTEMA DE DISTRIBUCIÓN ELÉCTRICA

Un sistema eléctrico de potencia (SEP) es un conjunto de centrales, líneas de transmisión y sistemas de distribución que operan como un todo. En operación normal todas las máquinas del sistema trabajan en paralelo y la frecuencia del SEP es constante [23].

La red de distribución es una parte muy importante del sistema eléctrico porque toda la electricidad producida debe distribuirse a los usuarios y esta electricidad se distribuye en un área amplia. Por otro lado, la producción de electricidad se concentra en grandes bloques de centrales eléctricas de gran capacidad y repartidas en grandes áreas con cargas de diferentes tamaños. Por lo tanto, el sistema de distribución de energía es aún más complejo que el sistema de energía [23].

Desde un punto de vista técnico, la definición clásica de un sistema de distribución de energía incluye lo siguiente;

1. Subestación principal,
2. Sistema de transmisión auxiliar,
3. Estación de distribución,
4. Alimentadores primarios,
5. Transformadores de distribución,
6. Servicios.

Estos elementos son aptos para cualquier tipo de carga en redes aéreas y subterráneas. La Figura 1 muestra los elementos principales del sistema de potencia y el sistema de distribución [23].

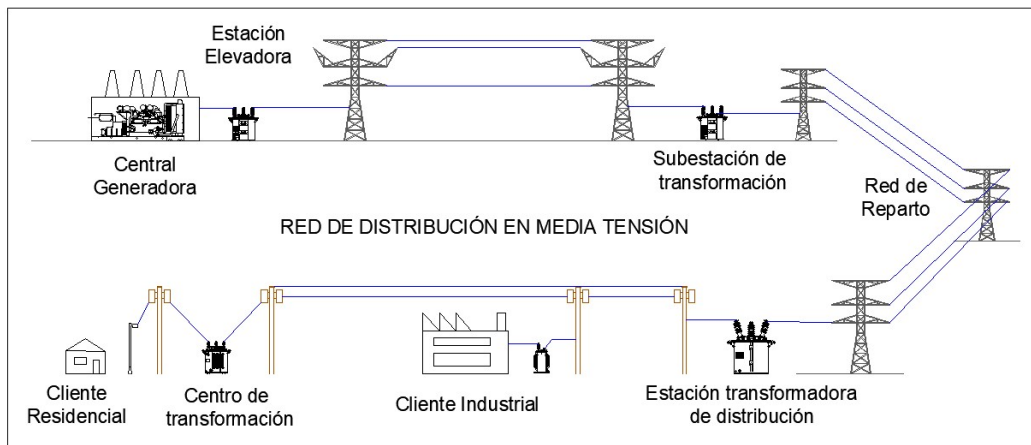


Figura 1. Configuración de un sistema eléctrico de potencia. Fuente:[23].

Las estructuras más comunes utilizadas en sistemas de subtransmisión, en alimentadores y redes secundarias son las siguientes:

1. Subtransmisión

- Radial
- Anillo
- Malla
- Anillo con amarre

2. Red Primaria

- Circulo alimentador primario radial
- Alimentador primario radial con enlace
- Alimentador primario con amarres de emergencia
- Malla primaria

3. Red Secundaria

- Secundario Radial
- Secundario en hilera
- Secundario en Malla
- Secundario Selectivo

La elección de la estructura del sistema de distribución de energía depende principalmente de la continuidad deseada de operación, regulación de voltaje y costo. El sistema a veces restringe la aplicación de ciertas estructuras, ya que el diseño se debe considerar las características del sistema, no las cosas individuales [23].

2.2 INTERRUPCIONES EN EL SISTEMA DE DISTRIBUCIÓN ELÉCTRICA

La sociedad actual depende de la disponibilidad constante de suministros eléctricos. En la mayoría de los países, el suministro de energía comercial se proporciona a través de redes nacionales que conectan múltiples plantas de energía a las cargas. La red debe satisfacer las necesidades básicas de iluminación, calefacción, refrigeración, aire acondicionado, transporte y vivienda de la nación, así como suministros esenciales para el gobierno, el público y las comunidades públicas, la industria, las finanzas, el comercio, la atención médica y los medios de comunicación. De hecho, el suministro de electricidad con fines comerciales es la base de un estilo de vida social. El tipo de maquinaria que implementa a nivel industrial, comercial o residencial que requiere un suministro constante o interrumpido. Por ejemplo, una falla de energía en una gran empresa financiera puede costar miles de dólares por minuto en tiempo de inactividad irreparable seguido de horas en tiempo de recuperación [24].

Debido a la expansión de la red de distribución, el sistema de alimentación ininterrumpida puede verse interrumpido por diversos eventos, como condiciones climáticas, fallas en los equipos, accidentes de tránsito u otras actividades relacionadas. Los problemas eléctricos que afectan a los equipos a menudo se generan localmente dentro de las instalaciones debido a diversas condiciones, como altas cargas de arranque, componentes de distribución de energía defectuosos e incluso ruido ambiental [25].

En caso de falla, las empresas de distribución decidieron minimizar la probabilidad de que se repita un incidente limitando la posibilidad de "peligrosidad" y aislando las partes sensibles del sistema. En un sistema distribuido (SD), muchas interrupciones son inevitables porque algunas de ellas son aleatorias [25].

2.3 INTERRUPCIONES

La discontinuidad se define como la pérdida total de voltaje o corriente. Las pausas se clasifican en temporales, temporales o permanentes en función de su duración.

Tabla 1. Clasificación de interrupciones. Fuente: [25].

Interrupciones	Duración
Instantánea	0.5 a 30 ciclos
Momentánea	30 ciclos a 2s
Temporal	2s a 2m
Sostenida	Mayor a 2m

Las causas de los cortes de energía pueden variar, pero generalmente se deben a algún tipo de daño en la red, como rayos, animales, árboles, accidentes automovilísticos, condiciones climáticas adversas (vientos fuertes, nieve o hielo) en la línea, etc.), falla de hardware o operación del disyuntor principal [25].

El término "fallo de energía" describe una situación en la red eléctrica en la que, debido a la naturaleza de la falla, los dispositivos de protección automáticos no pueden restaurar la energía y se requiere la intervención humana. Este término describe la situación con mayor precisión que el término comúnmente utilizado "interrupción". El término "interrumpido" en realidad se refiere a una condición en un sistema donde un componente ha dejado de funcionar como se esperaba. [26].

2.3.1 Interrupción preventiva

La interrupción ocurre cuando una pieza se deja intencionalmente sin usar durante un cierto período de tiempo, generalmente para la construcción, el mantenimiento preventivo o la reparación. [27].

2.3.2 Interrupción forzada

Interrupciones debido a la salida repentina de componentes, sistemas eléctricos no redundantes (radial), condiciones críticas para la estabilidad operativa (falla) o operación incorrecta del equipo (error humano). personas)[27].

2.3.3 Interrupción programada

Interrupción controlada consistente que obedece con los procedimientos de trabajo delegados planificados. Los usuarios afectados, son notificados inmediatamente [27].

2.3.4 Interrupción no programada

Las interrupciones forzadas o controladas no obedecen un horario de trabajo autorizado no se ajustan a ningún plan. Los usuarios afectados no fueron informados con antelación [27].

Cualquier interrupción o falla no planificada es causada por la naturaleza específica del circuito. Aunque no todas estas condiciones conducirán a la avería; esto no es una excepción si la protección del sistema está habilitada o el tiempo de falla es corto. Estas condiciones se pueden dividir en las siguientes categorías:

2.3.4.1 Cortocircuitos

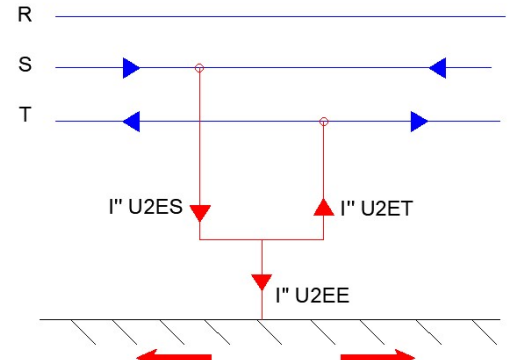
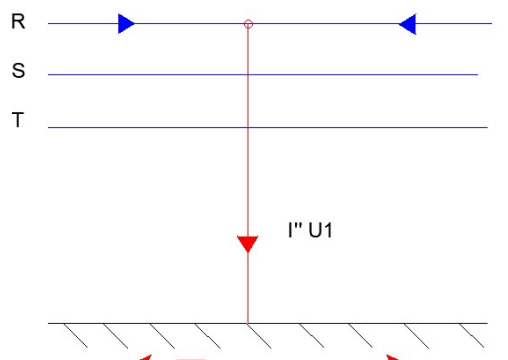
El cortocircuito o daño es una interferencia causada por una impedancia de contacto relativamente baja entre dos puntos de diferente potencial. Surgen de factores como la pérdida de aislamiento del hardware, eventos naturales o errores humanos. [28].

Según las etapas involucradas y la presencia de contactos de puesta a tierra, el cortocircuito puede ser monofásico, bifásico aislado, bifásico a tierra o trifásico. Los segundos errores se denominan errores de simetría porque implican el contacto físico de las tres fases, manteniendo así el sistema en equilibrio. Para estos errores, los errores residuales se denominan asimétricos porque provocan un desequilibrio del sistema. [2].

Los tipos de falta de pérdida total o parcial de tensión son los siguientes:

Tabla 2. Clasificación de los tipos de cortocircuitos. Fuente: [29].

TIPOS DE FALLAS	
<p>Fallas trifásicas</p> <p>Una falta trifásica es la única que se comporta como un circuito equilibrado, ya que afecta a todas las fases por igual. La tensión en el lugar de la falla, independientemente del cortocircuito a tierra o del aislamiento a tierra, es cero y la corriente tiene el mismo módulo, pero los parámetros se desplazan 120°. Este es uno de los casos de cómputo y cortocircuito más forzados. Como sistema balanceado, sus cálculos solo requieren el uso de una red secuencial directa [29].</p>	<p>a) Falla trifasica</p>
<p>Fallas bifásicas sin contacto a tierra</p> <p>En general, la corriente de cortocircuito simétrica inicial es menor que la corriente de cortocircuito trifásica, pero si el cortocircuito se produce cerca de un motor síncrono o asíncrono de cierta potencia, la corriente de cortocircuito puede ser aún mayor. en comparación con un cortocircuito trifásico. Dado que ocurre en dos de las tres fases del sistema, este cortocircuito ya no está balanceado, lo que obliga a su cálculo a utilizar redes de secuencia directa e inversa. [29].</p>	<p>b) Falla bifásica sin contacto a tierra</p>

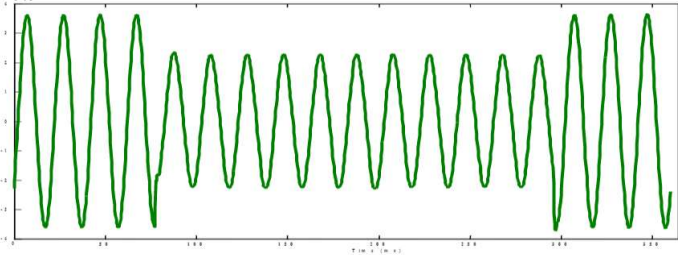
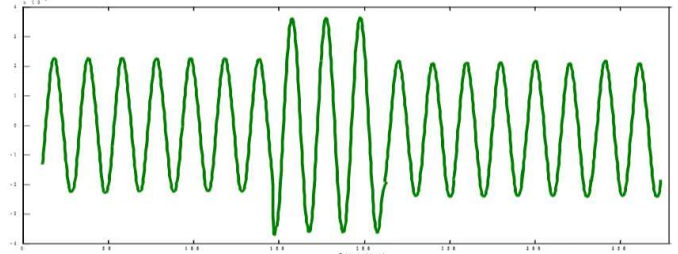
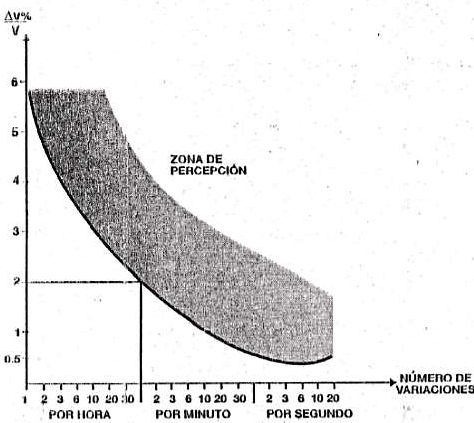
<p>Falla bifásica con contacto a tierra</p> <p>Tiene las mismas propiedades que un cortocircuito bifásico sin contacto a tierra, pero en este caso una pérdida de energía a tierra. En caso de fallo de este tipo, además de las redes de secuencia directa e inversa, también se debe tener en cuenta la red de secuencia copolar por pérdida de energía. [29].</p>	 <p>c) Falla bifásica con contacto a tierra</p>
<p>Falla monofásica</p> <p>Este es el cortocircuito más común y severo y ocurre con mayor frecuencia en redes con valores de baja impedancia. Su cálculo es importante porque tiene una corriente elevada y está conectado a tierra, lo que permite calcular fugas a tierra, tensión de contacto o tensión de paso, o evaluar las perturbaciones que estas corrientes pueden provocar. Para su cálculo se necesitan tres redes secuenciales (directa, inversa y unipolar) por desequilibrio y pérdida de energía. [29].</p>	 <p>d) Falla monofásica</p>

2.3.4.2 Sobretensión

Las sobretensiones transitorias son picos de tensión de corta duración que se producen en sistemas eléctricos o electrónicos, cuyo origen puede ser un rayo (directo o de corto alcance), grandes cargas capacitivas o inductivas de redes vecinas a un sistema eléctrico, una descarga electrostática o una explosión nuclear. De las primeras posibles causas, las emisiones a la atmósfera son las que presentan mayor riesgo por dos motivos principales: por un lado, por el nivel de tensión y corriente que generan, por otro, la exposición con riesgo no implica que los equipos electrónicos o los sistemas eléctricos entran en contacto físico; es suficiente interactuar a través de un posible acoplamiento electromagnético [30].

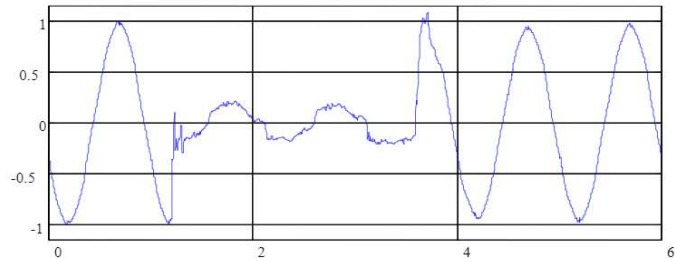
Los tipos de perturbaciones son las siguientes:

Tabla 3. Clasificación de los tipos de perturbaciones. Fuente:[31].

TIPOS DE PERTURBACIONES	
<p>Sags</p> <p>La norma CREG 024-2005 cita lo siguiente: “Un sag es una fluctuación de tensión caracterizada por producir una depresión transitoria de tensión respecto de la onda estándar, en un punto del SIN. Se considera un sag, el hundimiento con una duración de 0.5 ciclos a 1 min, con una magnitud igual o menor del 90% del valor rms de la onda” [32] [33].</p>	
<p>Swell</p> <p>La norma (CREG 024-2005) cita lo siguiente: “Un swell es la fluctuación de tensión caracterizada por producir un aumento transitorio de tensión respecto de la onda estándar, en un pulso del SIN. Se considera un swell a los picos de tensión entre 0.5 ciclos y 1 min, con una magnitud por encima al 10% del valor RMS de la onda” [32] [33].</p>	
<p>Flickers</p> <p>La norma (CREG 024-2005), con respecto a Flickers cita lo siguiente: “Flicker es la impresión de inestabilidad de la sensación visual causada por un estímulo luminoso, cuya luminosidad o distribución espectral fluctúa en el tiempo. El flicker es medido por un indicador de la perceptibilidad de un equipo o sistema, ante fluctuaciones de tensión durante un periodo de tiempo corto (10 minutos), obtenido de forma estadística a partir del tratamiento de la señal de tensión” [32] [34].</p>	

Armónicos

La norma (CREG 024 - 2005), recomienda acerca de THDV lo siguiente: “Es un indicador de la Distorsión Armónica Total de la Tensión, respecto de la onda estándar, expresada en porcentaje” [32] [33].

**2.3.4.3 Transitorios**

Los transitorios son quizás las perturbaciones energéticas más destructivas y se dividen en impulsos y oscilaciones [35].

- **Impulsivos**

Los transitorios de pulso son eventos pico repetitivos que aumentan los niveles de voltaje y/o corriente en una dirección positiva o negativa. Estos tipos de eventos se pueden clasificar en función de la rapidez con que ocurren.

Las causas de los transitorios incluyen rayos, conexión a tierra deficiente, cargas inductivas, rayos en la red y descargas eléctricas. Los efectos pueden incluir pérdida o corrupción de datos o daño físico al dispositivo [35].

- **Oscilatorios**

Los transitorios oscilantes son cambios abruptos de estado estable en el voltaje o la corriente de la señal, o los límites positivo y negativo de una señal, que oscilan a la frecuencia natural del sistema. El transitorio aumenta el voltaje a través de la señal de la fuente, seguido de una caída de voltaje muy brusca y alterna [26].

2.3.4.4 Pérdida de aislamiento

La pérdida de aislamiento es el deterioro real del material aislante de los equipos eléctricos. Es la causa principal de la corriente de fuga del conductor, la causa principal de la salida de energía del circuito normal. Este fallo puede provocar un cortocircuito en el sistema [36].

2.4 SISTEMA ELÉCTRICO CENTROSUR

CENTROSUR está ubicado en la provincia del Azuay, Cuenca. Es la empresa encargada de la distribución de energía eléctrica a la mayor parte del Azuay, Cañar y Morona Santiago, cubre el 11,79% del territorio del Ecuador [37]. Tiene 30.243 km^2 , divididas en zonas (zonas 1, 2, 3) y una vía troncal que se contempla desde 2014. La mayor parte del área de franquicia de la empresa es la provincia de Morona Santiago con un margen de utilidad del 62,84% como porcentaje, seguida por Azuay, Cañar y otras provincias (Naranjal, El Guabo, Saraguro y El Piedrero) con 0,78% [37].

2.4.1 ANÁLISIS DE DATOS DE LA CENTROSUR

En el Anexo A proporciona datos históricos sobre las interrupciones de la Empresa Eléctrica (Interrupciones no programadas). Se presenta una clasificación de estas interrupciones que específicamente la Empresa Eléctrica divide según el origen en red, ambientales, climáticas, maniobras, mantenimiento, u otras,). Se propone categorizar estas exenciones por medios en función del origen de la red, entorno, clima, maniobras, mantenimiento u otros detalles). De acuerdo con la información brindada, en el período 2015-2019 se utilizó una base de datos de 13,551 cortes, en el caso de los interruptores, el mayor porcentaje de cortes en los años 2015, 2016 y 2017 se debió a cortocircuitos, mientras que en 2018 y 2019, la tasa de cortes fue la más alta. tasa de tiempo de inactividad debido a pausas temporales.

Las fallas de energía por cortocircuitos y fallas transitorias son comunes en las empresas de distribución, especialmente en la CENTROSUR.

2.5 NORMAS EN EL SERVICIO ELÉCTRICO DE DISTRIBUCIÓN

La empresa es responsable de desarrollar, a través de su División de Sistemas Analíticos y de Distribución (SIGADE), la información técnica y de gestión necesaria para medir el nivel de calidad de la energía entregada en las distintas etapas de operación.

En cuanto al almacenamiento de información, CENTROSUR cuenta con una base de datos de interrupciones planificadas y no planificadas desde el año 2008, lo cual es un requisito de la Regulación No. CONELEC – 004/01. La información de cada interrupción se registra de la siguiente manera:

- Zona: Se registra en una de las tres áreas donde se han informado interrupción en el alimentador.
- Alimentador: Se enfatiza al número del alimentador que fue afectado y dañado.
- kVA instalado: son los KVA instalado en el mes de observación.
- Causas y sub causas: Se identifican causas externas, internas, fuera de plan, y cuando se trata de causas secundarias tenemos diferentes causas como clima, emisiones atmosféricas, desconocido, equipo, vida silvestre, otros, planeado, público y vegetación.
- Año: Cada año se actualizan todos los errores que surgen en el sistema.
- Mes de visualización: El sistema no funcionará sin problemas, por lo que el sistema se monitorea diariamente de enero a diciembre de cada año, y si el sistema SCADA detecta un error, el SIGADE registrará el mes que provocó la falla.
- Nodo: Registra el número de dispositivos de protección activados en caso de falla.
- Número de Interrupción: Recuento de errores de registro actualizado diariamente.
- Fecha de desconexión: Iniciaré sesión en el año, mes, día y hora en que ocurrió el problema.
- Potencia en kVA: Es la potencia en kVA de la red desconectada.
- Falla de Origen: Se captura a partir de interpretaciones generadas por la EERCS e incluye tres métricas por las que afecta el sistema (bajo, medio, alto.) tensión, el tipo de daño (ramal, general y puntual) y su origen (planeado o no planeado) - Programado y Externo.
- Tiempo en horas: Tiempo (en horas) para registrar el retraso de desconexión.

- Origen: Registra el tiempo de inactividad planificado (mantenimiento, etc.) o no planificado (rayos, vegetación, etc.).
- Dirección: Registrar las avenidas y calles desde donde se realizó el punto de daño.
- Coordenadas en X: Esta es la ubicación del dispositivo más cercano al error, expresada en longitud.
- Coordenadas en Y: Es la ubicación del dispositivo más cercano al error, expresada en latitud.
- FMIK: Esta es la frecuencia promedio de la interrupción (los kVA promedio que han sido desgastados por una interrupción.).
- FMIKZ: Se refiere a la frecuencia promedio de interrupciones en un área determinada.
- FMKA: Se refiere a la frecuencia de interrupción promedio del alimentador.
- TTMIK: Indica el tiempo de inactividad total del sistema (tiempo total durante el cual no se utiliza el promedio de kVA).
- TTMIKZ: El tiempo de inactividad total para esa área.
- TTMKA: Se refiere al tiempo total de inactividad del alimentador.
- Clientes: Número de clientes afectados por cada falla.
- Subsistema: Registrar a qué tensión se produce la avería (baja, media o alta tensión).

2.6 DISPOSICIONES DE REGULACIONES VIGENTES

Las empresas de energía en Ecuador deberán cumplir con la normativa publicada por ARCONEL, a partir del 16 de enero de 2016, para mejorar el servicio bajo la Ley de Servicio Público [38].

La Regulación No.004/01 de calidad de energía, dice que “Se consideran todas las interrupciones mayores a 3 minutos, incluyendo las de origen externo, debidas a fallas en transmisión. No se consideran las interrupciones con duración igual o menor a 3 minutos” [38]. La misma Regulación también estipula lo siguiente: “No se considera interrupciones de un consumidor en particular causadas en sus propias instalaciones, siempre y cuando no afecten a otros consumidores” [38].

A modo de resumen, las actividades que debe cumplir una empresa distribuidora bajo el CONELEC - 004/01 son: [38]:

1. Los distribuidores deben mantener registros históricos de al menos los últimos tres años.
2. El acta deberá contener la fecha, lugar y hora de inicio, inicio y finalización de la interrupción.
3. El distribuidor debe presentar un informe mensual a la unidad especificada. El ceñirse a estos tres puntos permite obtener datos de interés para este trabajo.

III CAPÍTULO

APM Y HERRAMIENTAS DIGITALES

Este capítulo detalla diferentes métodos de predicción y las herramientas de software de geoprocésamiento como ArcGIS®, Python™ son herramientas clave en el proceso de simulación.

3.1 INTRODUCCIÓN

El cerebro humano tiene una arquitectura compleja de neuronas biológicas que están interconectadas para formar una extensa red capaz de transmitir impulsos neuronales a través de las sinapsis (las áreas de comunicación entre los extremos de una neurona y otra). Esta red transforma las reacciones de los elementos electroquímicos en conceptos mentales complejos como el aprendizaje, la memorización y el razonamiento [22].

Durante la última década, se han desarrollado procesadores de computadora con un poder de cómputo extraordinario. Los cálculos matemáticos largos y las clasificaciones complejas son tareas que las computadoras pueden realizar fácilmente en comparación con el tiempo que les toma a los humanos realizar estas tareas. Deshabilitar o activar un procesador solo toma de 3 a 10 segundos, pero hay tareas que los humanos pueden realizar de manera más eficiente que las computadoras, como: B. Reconocer frutas y distinguir los colores de las herramientas. Estos procesadores facilitan la realización de cálculos matemáticos dentro de una red neuronal, independientemente de la cantidad de cálculos que deban realizarse [22].

3.2 REDES NEURONALES.

Las neuronas reciben señales (entradas) de otras neuronas vía conexiones sinápticas que pueden ser excitantes o inhibitoras. En función de las señales recibidas, una neurona envía a su vez una señal a otras neuronas por medio del axón en la figura 2 se muestra una red neuronal biológica [39].

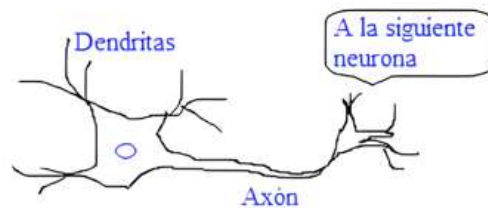


Figura 2. Neurona biológica. Fuente: [39].

Las neuronas son importantes unidades de procesamiento de información en las redes neuronales. Con base en esto, se utilizan métodos para predecir las necesidades. Se parece a un sistema nervioso biológico formado por neuronas que funcionan en paralelo y se puede adaptar al proceso de aprendizaje de almacenamiento de conocimientos. Esto es lo que él también llama una máquina adaptativa. Las redes neuronales artificiales se pueden definir como estructuras de procesamiento de información distribuidas en paralelo que consisten en neuronas interconectadas por conexiones sinápticas. Cada neurona tiene una única salida distribuida a través de múltiples conexiones bilaterales [40].

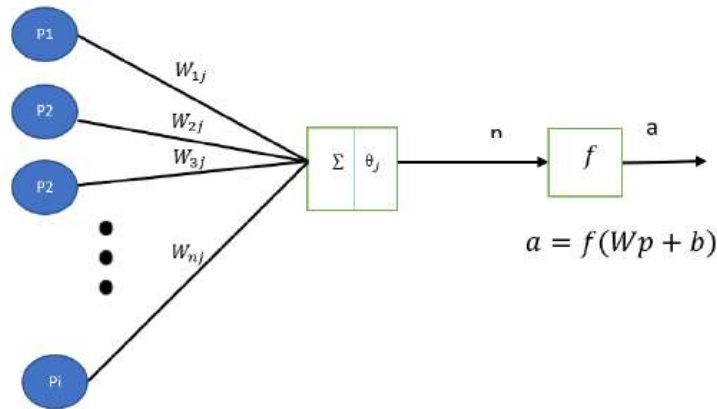


Figura 2. Ejemplo básico de una estructura neuronal. Fuente:[40].

En la Figura 3 se muestra una representación esquemática del modelo matemático de una neurona. Se puede expresar una relación entre las entradas y salidas de la red neuronal. En la Ecuación 1 se representa como:

$$a = f(\sum_{i=1}^n P_j) * (W_{ij} - \theta_j) \text{ Ecuación 1}$$

Donde:

P_j ($i=1,2, 3, \dots, n$) = señales de ingreso de n neuronas externas a una neurona interna.

W_{ij} = equivale al peso asociado a cada una de las conexiones entre dos neuronas.

θ_j = umbral de activación de la neurona j .

f = es la función de activación (signo o sigmoidea)

a = es la salida de la red neuronal.

3.2.1 FUNDAMENTOS

Se puede entender a una red neuronal como un arreglo de capas de neuronas artificiales, distribuidas de tal manera que formen una estructura capaz de realizar los cálculos para la cual fue diseñada. Un conjunto de neuronas artificiales forma una capa de neuronas, así mismo, un conjunto de capas forma una red neuronal [22].

Cada unidad o neurona artificial calcula su salida en función de la entrada que recibe. La salida de cada unidad puede servir como entrada de otra unidad. La red toma un conjunto de entradas externas y devuelve la salida de algunas de sus neuronas al mundo exterior. Los cálculos matemáticos realizados en cada unidad no son muy complicados [22]. En la Figura 4 se observa cómo está estructurada una red neuronal.

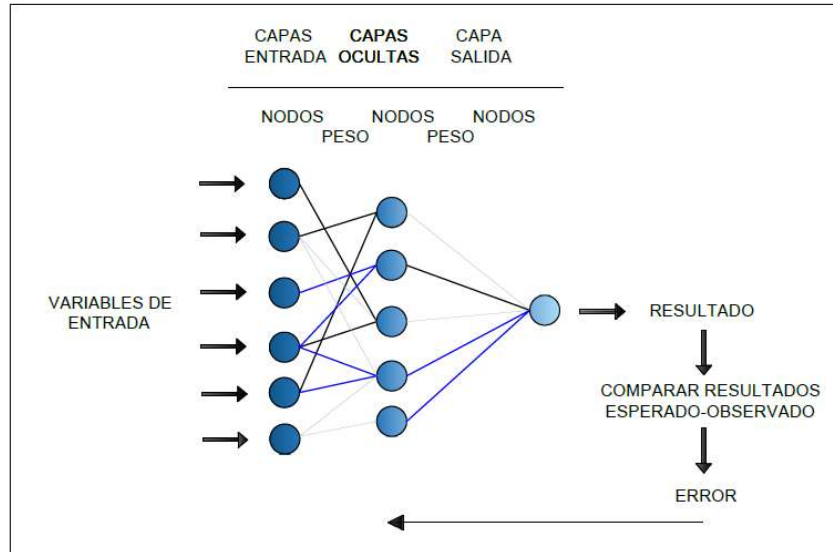


Figura 3. Secciones de una red neuronal.

En las neuronas artificiales, cada arco tiene un peso numérico asociado, indicado por w_{ij} , que determina la fuerza y el signo de la conexión (imita una sinapsis). La figura 5 muestra los pesos numéricos de las neuronas.

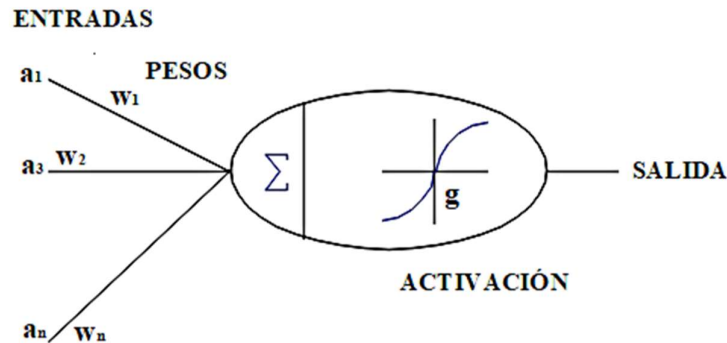


Figura 4. Pesos numéricos de una neurona artificial [22].

La salida y_i de una neurona se calcula como indica en la Ecuación 2:

$$y_i = g\left(\sum_{j=0}^n w_{ij} a_{ij}\right) \text{ Ecuación 2}$$

Donde:

- g es una función de activación.
- La sumatoria $\left(\sum_{j=0}^n w_{ij} a_{ij}\right)$ se hace sobre todas las unidades j que envían su salida a la unidad i , esta se conoce como regla de propagación.
- w_{ij} : de cada unidad se interpreta como una cantidad que debe superar la suma de las señales de entrada que recibe la unidad, para que se active.
- La función de activación g tiene el papel de normalizar la salida cuando el umbral de entrada se supera. Además, hace que la red no se comporte simplemente como una función lineal.

- Si los índices i y j comienzan en 0 y se denota la entrada $a=x$, el peso umbral (bias) $w_{i0}=\theta_i$ y $x_0=-1$, podemos expresar en la Ecuación 3:

$$y_i = g\left(\sum_{j=1}^n w_{ij}x_j - \theta_i\right) \text{ Ecuación 3}$$

Existen varias funciones de activación que se usan para diferentes propósitos, siendo un criterio de selección el tipo de objetivo que se persiga; entre las funciones de activación más conocidas están las siguientes:

- **Función bipolar (sgn)**

La función bipolar se calcula como se muestra en la Ecuación 4, y en la Figura 6 se representa gráficamente.

$$\text{sgn}(x) = \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x \leq 0 \end{cases} \text{ Ecuación 4}$$

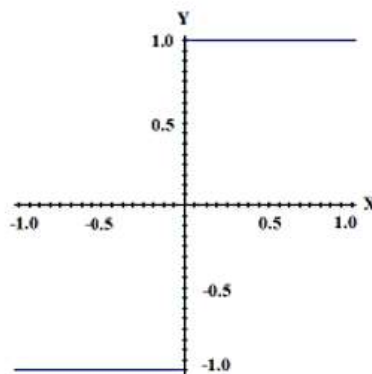


Figura 6. Grafica de función bipolar [22].

- **Función tangente hiperbólica (tanh)**

La función tangente se calcula como se muestra en la Ecuación 5, y en la Figura 7 se representa gráficamente.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \text{ Ecuación 4}$$

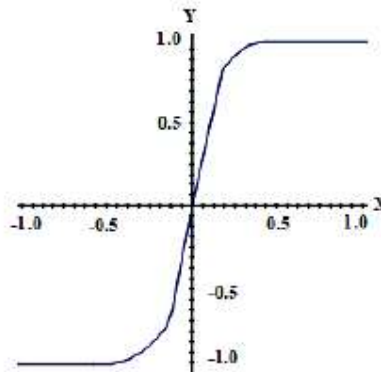


Figura 5. Grafica de tangente hiperbólica [22].

- **Función sigmoide (σ)**

La función sigmoide se calcula como se muestra en la Ecuación 6, y en la Figura 8 se representa gráficamente.

$$\sigma(x) = \frac{1}{1+e^{-x}} \text{ Ecuación 6}$$

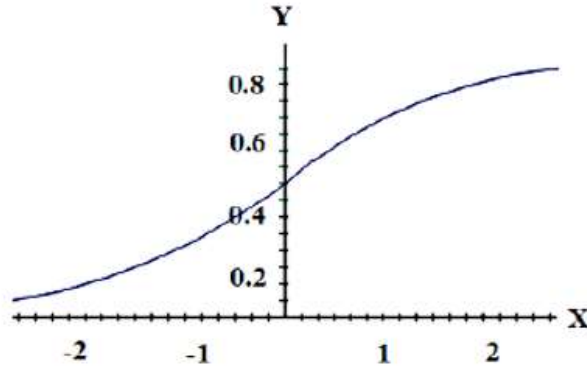


Figura 8. Grafica de función sigmoide [22].

- **Función unidad lineal rectificada (Relú)**

La función unidad lineal rectificada se calcula como se muestra en la Ecuación 7, y en la Figura 9 se representa gráficamente.

$$x^+ = \max(0, x) \text{ Ecuación 5}$$

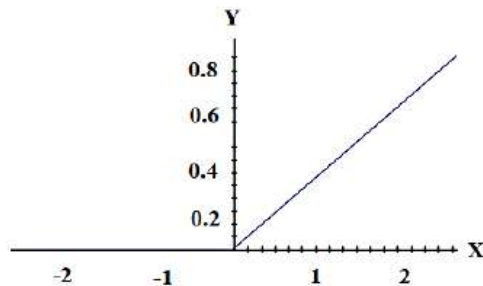


Figura 6. Grafica función Relú [22].

Por su uso en APM, la popularidad de la función Relú se ha incrementado; usualmente es usada en las capas de neuronas ocultas y demuestra algunas desventajas en procesos de salida por su característica de rampa.

3.1.1 ESTRUCTURA DE REDES NEURONALES.

En la Figura 10, se define los componentes de una red neuronal, se muestra una representación gráfica de la arquitectura de ANN conformada por tres capas: entrada, capas ocultas y capa de salida.

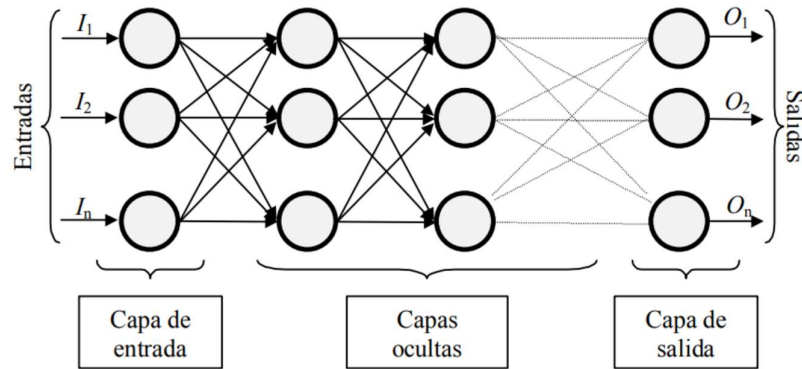


Figura 7. Ejemplo de una red neuronal. Fuente:[39].

Existe diversidad de estructuras que pueden ser usadas para las redes, dependiendo principalmente de su aplicación, pero cada una de estas sigue el siguiente proceso [39]:

- Las redes neuronales artificiales están compuestas de gran cantidad de procesadores conectados entre sí y actuando en paralelo. Los modelos neuronales biológicos son mucho más complejos que los modelos computacionales actuales.
- El comportamiento de la red está determinado por su topología, los pesos de las conexiones y la función característica de las neuronas.

3.1.2 CARACTERÍSTICAS DE LAS REDES NEURONALES.

El comportamiento de una red está determinado por su topología, los pesos de las conexiones y la función característica de las neuronas [39]. Para ello existen varias características que una red neuronal debe cumplir como se muestra a continuación:

- **Aprendizaje Adaptativo**

Las ANN aprenden a realizar tareas a partir de un conjunto de datos dados en el proceso de aprendizaje [39].

- **Autoorganización**

Pueden crear su propia organización o representación de la información recibida [39].

- **Operación en tiempo real**

Las operaciones realizadas pueden ser llevadas a cabo por computadores paralelos, o dispositivos de hardware especiales que aprovechan esta capacidad [39].

- **Tolerancia a fallos parciales**

La destrucción parcial de una red daña parcialmente el funcionamiento de la misma, pero no la destruye completamente. Esto es debido a la redundancia de la información contenida [39].

3.1.3 CLASIFICACIÓN DE LAS REDES NEURONALES

Los elementos ya mencionados se pueden conectar para dar forma a modelos o estructuras neuronales diferentes, pudiendo ser clasificados por:

1. El número de capas.
2. Forma de conexión.

3. Forma de asociación.
4. Dinámica de actualización.

Durante esta fase, la red es responsable de mantener el peso de la conexión y ajustar la red a objetivos específicos. La experiencia de las redes neuronales se puede expresar aprendiendo y cambiando pesos y conexiones sinápticas, incluso si hay dos o más redes idénticas que comparten la misma estructura que la entrada, con resultados diferentes [40].

1. Criterio de parada

El criterio de parada dentro de este tipo de algoritmos es importante debido que se reduce el tiempo de aprendizaje y mejora la eficiencia de la red optimizando la predicción [40].

Existen dos fases en toda aplicación de las redes neuronales: la fase de aprendizaje o entrenamiento y la fase de prueba [39].

a. Fase de Aprendizaje

Una característica de las redes neuronales es su capacidad de aprender. Aprenden por la actualización o cambio de los pesos sinápticos que caracterizan a las conexiones. Los pesos son adaptados de acuerdo a la información extraída de los patrones de entrenamiento nuevos que se van presentando. Normalmente, los pesos óptimos se obtienen (minimizando o maximizando) alguna "función de energía". Por ejemplo, un criterio popular en el entrenamiento supervisado es minimizar el least-square-error (error cuadrático medio) entre el valor deseado y el valor de salida de la red [39].

3.1.4 ARQUITECTURA

Hace referencia al patrón o forma de conexión de las neuronas, entre algunos parámetros tenemos: número de capas, número de neuronas por capa y forma de conexión, para esta última podemos dividirla en dos [40]:

- Redes feed-forward (hacia adelante) que conecta las neuronas en una sola dirección hacia capas posteriores.
- Redes recurrentes (hacia atrás) que utilizan retroalimentación consigo mismas o con otras neuronas de una capa anterior, formando un bucle.

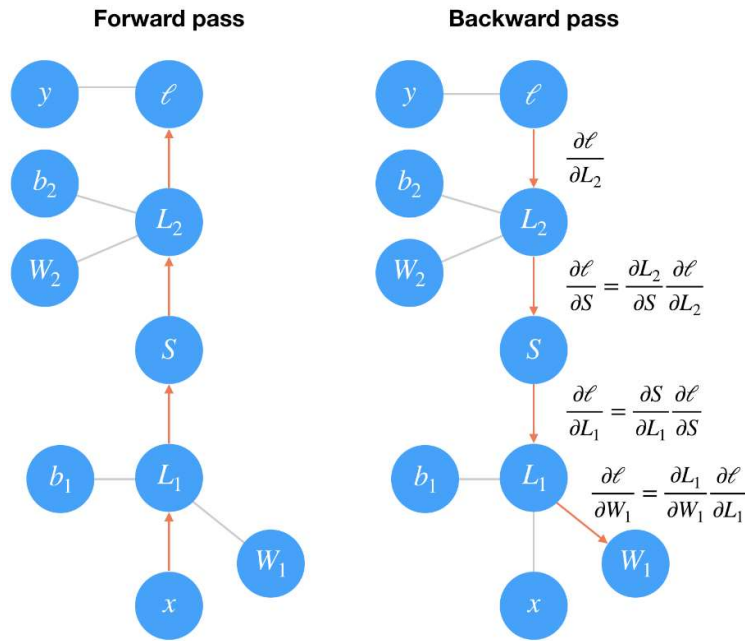


Figura 8. Camino de la red para el entrenamiento Feedwars y Backward. Fuente: [41].

En el entrenamiento para adelante o forward pass las operaciones van de abajo hacia arriba, como se observa en la Figura 11 mediante la siguiente serie de pasos:

- **Paso 1:** las entradas x pasan por una transformación lineal $L1$ con pesos $W1$ más las bias $b1$.
- **Paso 2:** La salida pasa por la función de activación (sigmoideal) S y se someten a otra transformación lineal $L2$ con nuevos pesos $W2$ y nuevas bias $b2$.
- **Paso 3:** Finalmente se calculan las pérdidas l , que presentan una medida de cuan malas son las predicciones encontradas por la red.

Mientras que para el entrenamiento hacia atrás (backward pass) utiliza el descenso de gradiente, es la propagación hacia atrás a lo largo de la red. Cada operación tiene cierto gradiente entre las entradas y las salidas. Conforme enviamos los gradientes hacia atrás, multiplicamos el gradiente entrante con el gradiente de la operación. Matemáticamente se define como la aplicación de la regla de la cadena, calculando el gradiente de perdida con respecto a los pesos como indica la Ecuación 8 [41].

$$\frac{\partial \ell}{\partial w_1} = \frac{\partial L_1}{\partial w_1} \frac{\partial S}{\partial L_1} \frac{\partial L_2}{\partial S} \frac{\partial \ell}{\partial L_2} \text{ Ecuación 6}$$

Se actualizan los pesos usando el gradiente con una tasa de aprendizaje α como se muestra en la Ecuación 9.

$$w'_1 = w_1 - \alpha \frac{\partial \ell}{\partial w_1} \text{ Ecuación 7}$$

El valor de α se establece de manera que los pasos de actualización de peso sean lo suficientemente pequeños como para que el método iterativo se establezca en un mínimo. Las redes neuronales tienen, según la forma de conexión y propiedades, diferentes tipos de arquitectura. El fundamento matemático de las redes neuronales es extenso; no es el objetivo de este trabajo su estudio en detalle, sin embargo, se hará énfasis en los conceptos de vital importancia para la comprensión de cómo resolver el problema a tratar mediante el uso de una RN [22].

3.2.5.1 Perceptrón simple

El concepto introducido en [42], muestra un sistema unidireccional compuesto por n entradas y m salidas. La operación para el cálculo de propagación de salida y_i de un perceptrón simple se expresa como en la Ecuación 10:

$$y_i = g(\sum_{j=1}^n w_{ij}x_j - \theta_i) \text{ con } i = 1, \dots, m \text{ Ecuación 8}$$

Los autores en [42] señalan que, con 2 entradas, x_1 y x_2 ubicadas en las abscisas y ordenadas respectivamente, se tiene la Ecuación 11:

$$y = \begin{cases} 1 & \text{si } w_1x_1 + w_2x_2 \geq \theta \\ 0 & \text{si } w_1x_1 + w_2x_2 < \theta \end{cases} \text{ Ecuación 9}$$

Trabajando sobre x_2 se tiene la Ecuación 12 y 13, y en la Figura 12 se muestra la región de un perceptrón simple.

$$w_1x_1 + w_2x_2 - \theta = 0 \text{ Ecuación 10}$$

$$x_2 = \frac{w_1}{w_2}x_1 + \frac{\theta}{w_2} \text{ Ecuación 11}$$

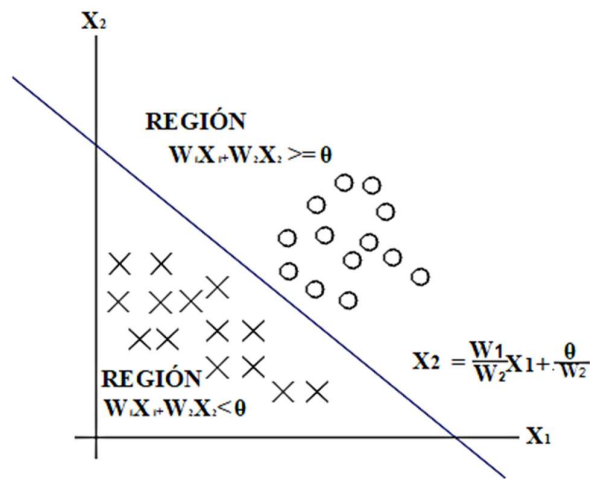


Figura 9. Región de decisión de un perceptrón simple. Fuente: [22] .

Este perceptrón se entiende como un clasificador. El algoritmo de Rosenblatt determina automáticamente los pesos sinápticos que clasifican patrones a partir de ejemplos etiquetados. Es conocido también como un discriminador lineal que ajusta el peso dependiendo del número t de iteraciones y los N pares de patrones entrada-salida según la Ecuación 14:

$$w_{ij}(t + 1) = w_{ij}(t) + \sum_{r=1}^N \Delta w_{ij}^r(t) \text{ con } r = 1, \dots, N \text{ Ecuación 12}$$

Un perceptrón simple como discriminador puede clasificar datos dividiendo regiones mediante un hiper-plano o un plano como se observa en la Figura 12. Pero para tareas más complejas como divisiones no lineales existen herramientas que se han desarrollado tomando en cuenta este tipo de limitaciones.

3.2.5.2 Perceptrón multicapa.

El perceptrón multicapa (PMC) introduce capas ocultas en una red, esto hace que varios de los problemas del perceptrón simple sean solucionados. Estos algoritmos usan para su entrenamiento un algoritmo de retro propagación para actualizar sus pesos, de ahí que esta arquitectura se conozca también como redes de retro propagación.

El autor en [43] introdujo por primera vez en su tesis doctoral la teoría del PMC, pero no fue hasta 1986 que Rumelhart lo popularizó gracias a que la computación había avanzado al punto de satisfacer los requisitos que el algoritmo de retro propagación demandaba [43].

Considerando una capa oculta con función de activación sigmoide y una capa de salida con función de activación lineal, la salida z_k de la capa final de cada neurona k en una red PMC se calcula como indica la Ecuación 15:

$$z_k = \sum_{j=1}^0 w'_{kj} y_j - \theta'_k = \sum_{j=1}^0 w'_{kj} f(\sum_{i=1}^n w'_{ki} x_j - \theta_j) - \theta'_k \text{ Ecuación 13}$$

Los diferentes θ_j representarán los umbrales correspondientes a la capa oculta, w'_{kj} los pesos de la capa de salida y θ'_k sus umbrales respectivos.

Dado un patrón de entrada x^r ($r = 1, \dots, N$) se puede generalizar la Ecuación 16 para cualquier número k de neuronas de salida:

$$z^r_k = \sum_{j=1}^0 w'_{kj} y_j - \theta'_k = \sum_{j=1}^0 w'_{kj} f(\sum_{i=1}^n w'_{ki} x_j^r - \theta_j) - \theta'_k \text{ Ecuación 14}$$

Adicional a esto, la arquitectura PMC, adiciona 2 procesos, el primero consiste en la regla de aprendizaje, usando la regla de los mínimos cuadrados o LSM (Least Mean Square) y un optimizador basado en la regla LSM conocido como el descenso por el gradiente. El gradiente $\nabla E(w)$ a su vez es la derivada parcial del error con respecto al peso y es usado para encontrar una configuración de pesos óptima hasta que el error sea mínimo. Matemáticamente se recorre la hiper-superficie del error hasta encontrar un mínimo local, el proceso es iterativo y se expresa como se muestra en la Ecuación 17:

$$w(t+1) = w(t) - \eta \nabla E(w) \text{ Ecuación 15}$$

Para iniciar el algoritmo se elige un valor aleatorio de pesos y umbrales iniciales ($t = 0$). La constante η es positiva y pequeña, llamada factor de aprendizaje, modera la actualización de pesos. Se debe iniciar con pesos pequeños, positivos y negativos, se debe elegir también la variación de pesos para cada conjunto de entrenamiento, este proceso se conoce como aprendizaje por lotes.

El algoritmo de retro propagación se puede justificar formalmente como un descenso por el gradiente del error, la salida de cada unidad se calcula propagando valores hacia adelante, pero el error en cada una se calcula desde la capa de salida hacia atrás (de ahí el nombre de retro propagación) [44].

Un conjunto de entrenamiento D debe tener ejemplos de la forma $(x^{\vec{}}, y^{\vec{}})$, con $x^{\vec{}} \in Rn$ y $y^{\vec{}} \in Rm$ (n y m), dimensión de la entrada y salida respectivamente), un factor de aprendizaje η , una función de activación g diferenciable y una estructura de red.

En la Figura 13, se muestra esta arquitectura. Se advierten que se puede realizar separaciones complejas a diferencia de un perceptrón simple, por ejemplo, la función XOR, ya que a diferencia de la función AND y OR, no es linealmente separable.

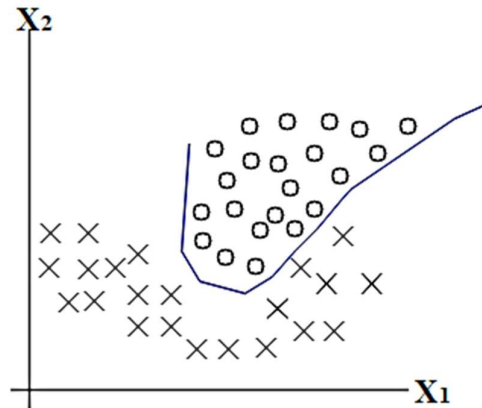


Figura 10. Región de decisión de la RN con arquitectura PMC. Fuente: [22].

3.2.5.3 Redes neuronales recurrentes.

Si bien las redes PMC pueden discriminar situaciones no lineales, existen tareas dificultosas para esta arquitectura, como permitir que una salida alimente a las entradas. La conexión completa y recurrente de la red, los algoritmos que forman una red neuronal recurrente (RNR) satisfacen muchas de las problemáticas presentes en arquitecturas más simples como en la Figura 14. El propósito de una RNR, es otorgar una suerte de memoria a la red.

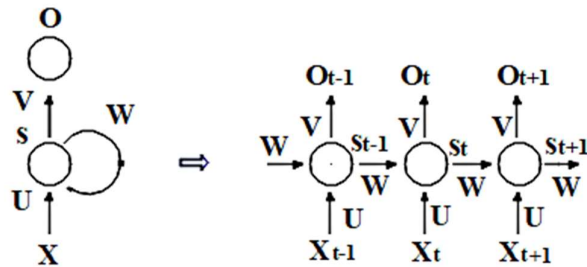


Figura 11. RNR y despliegue. Fuente: [22].

- x_t es la entrada en el tiempo t , por ejemplo, x_1 puede ser un vector correspondiente a la segunda palabra de una oración.
- s_t es el estado oculto en el tiempo t es calculado en base a los estados ocultos previos y en la entrada en el justo momento: $s(t)=(Ux_t-Ws_{t-1})$.
- o_t es la salida en el tiempo t . Por ejemplo, si queremos predecir la siguiente palabra de una oración, será un vector de probabilidades a través de un conjunto de entrenamiento. $o_t=softmax(Ust)$ [45].

Una RNR debe ser debidamente diseñada y entrenada antes de poder ejecutarse. El entrenamiento se da sobre un conjunto de datos referentes al problema. Por su parte, el diseño considera una

amplia selección de hiper-parámetros tales como: el porcentaje de datos usados para el entrenamiento, porcentaje de datos usados para la validación, función de activación, número de iteraciones o épocas, aprendizaje por lotes, etc.

3.2.5.4 Redes de memoria a largo plazo.

Las RMLP son muy populares en la actualidad por sus aplicaciones en pronósticos y reconocimientos, no poseen una diferencia fundamental de las RNR, pero se distinguen en los cálculos que realizan para el almacenamiento de memoria en la red.

La memoria en las RMLP puede entenderse como una caja negra donde se procesa el estado oculto anterior (h_{t-1}) y la entrada actual (x_t) de una neurona, cargando la responsabilidad de elegir lo que se conservará para la propagación de datos o lo que se eliminará en el paso a la siguiente capa de neuronas.

En [46] se expone el cálculo para una red RMLP de una manera completa, pero sin incurrir en exhaustiva complejidad. Define el área de toma de decisiones con respecto a la memoria como celdas con un respectivo estado.

La primera etapa del cálculo como indica la Ecuación 18 consiste en decidir que se suprime en el estado de celda, la decisión se realiza mediante una capa sigmoide llamada puerto de olvido, que analiza (h_{t-1}) y (x_t), con la adición de un nuevo umbral (b_f) y produce una salida entre 0 y 1 para cada número en el estado de celda C_{t-1} . El 0 significa olvidar completamente la característica, y el 1 mantener completamente la característica.

$$f(t) = \sigma(w_t[h_{t-1}, x_t] + b_f) \text{ Ecuación 16}$$

En la segunda etapa Ecuación 19 y 20, se decide lo que se va a almacenar en la celda, primero una capa de función σ , llamada puerto de entrada, toma en cuenta solo los valores se van a actualizar, para que una capa con función \tanh cree un vector con valores nuevos \bar{C}_t que serán añadidos al estado actual de la celda

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \text{ Ecuación 17}$$

$$\bar{C}_t = \tanh(w_c[h_{t-1}, x_t] + b_c) \text{ Ecuación 18}$$

Ahora que los dos puertos, el de entrada y el de olvido, han tomado las decisiones de lo que se va a conservar y de lo que no, se deben cumplir estas decisiones. Actualizando el anterior estado de la celda C_{t-1} a uno nuevo C_t , se multiplica el antiguo estado por $f(t)$ (olvidando lo que se decidió olvidar) se añade $i_t * \bar{C}_t$. Estos son los nuevos valores que se muestra en la Ecuación 21.

$$C_t = f_t(C_{t-1} + i_t)(\bar{C}_t) \text{ Ecuación 19}$$

Por último, en la Ecuación 22 y 23, se decide cuál va a ser la salida, esta decisión estará gobernada por el estado de la celda, se ejecuta una capa sigmoide que considera que parte de la celda va a mostrarse en la salida. Luego una función \tanh coloca a los valores en un rango entre -1 y 1, y después lo multiplica por la salida de la capa sigmoide, así solo se muestra como salida la parte deseada.

$$O_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \text{ Ecuación 20}$$

$$h_t = O_t * \tanh(C_t) \text{ Ecuación 21}$$

Por la eficacia mostrada para guardar valores relevantes y persistencia en la memoria, las RMLP han sido usadas ampliamente para realizar predicciones y reconocimientos de todo tipo.

3.1.5 CLASIFICACIÓN DE LOS MÉTODOS DE INTELIGENCIA ARTIFICIAL

Existen diferentes criterios para la clasificación de métodos de inteligencia artificial (IA) como los que se muestra en la Figura 15:

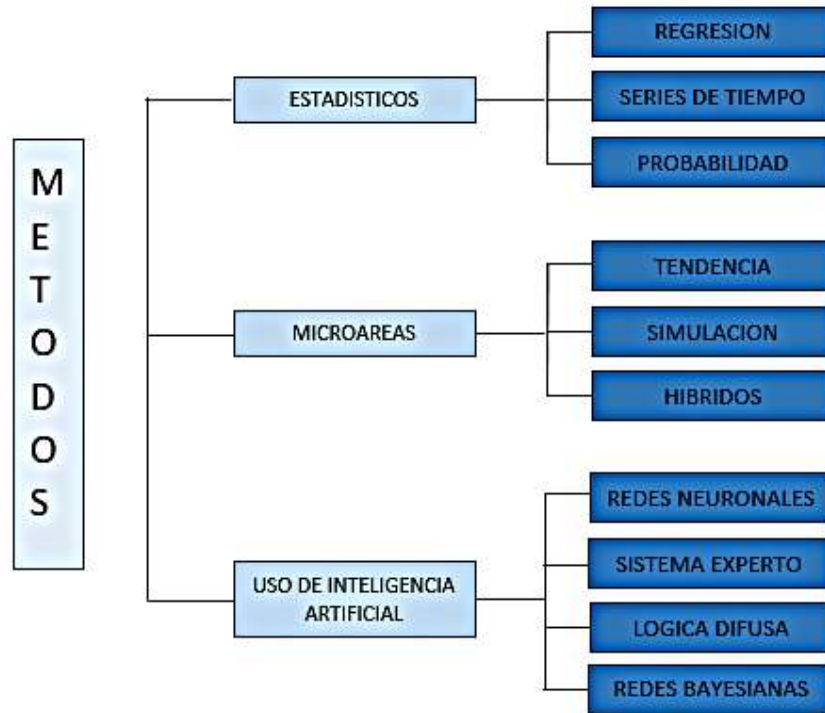


Figura 12. Clasificación de métodos. Fuente: [40].

3.1.6 USO DE LA INTELIGENCIA ARTIFICIAL

La inteligencia artificial (IA) hace referencia a la posibilidad de simular funciones cognitivas de los procesos mentales, como es el procesamiento, aprendizaje de información y/o resolución de problemas. IA ha desarrollado modelos para reconocimiento de voz, automatización, optimización, toma de decisiones, plataformas computacionales *“Machine Learning”*.

3.2.7.1 CARACTERÍSTICAS DE LOS MODELOS DE LA IA.

Los modelos y algoritmos “estándar” de la IA tienen las siguientes características:

- El conocimiento se representa explícitamente usando reglas, redes semánticas, modelos probabilísticos, etc.,
- Se imita el proceso humano de razonamiento lógico para resolver los problemas, centrandó la atención en las causas que intervienen en el problema y en sus relaciones (encadenamiento de reglas, inferencia probabilística).
- Se procesa la información secuencialmente.

Con el rápido desarrollo de la IA aparecieron un gran número de problemas complejos donde no se disponía de una representación explícita del conocimiento y de un procedimiento de razonamiento lógico para resolverlo. Posibles causas de este fallo: Procedimientos algorítmicos y estructura computacional empleados eran incorrectos [39].

3.2.7.2 APRENDIZAJE AUTOMÁTICO.

Un programa de computadora aprende de la experiencia (E) con respecto a alguna clase de tareas (T) y medida de desempeño (P), si su desempeño en tareas en (T), medido por (P), mejora con la experiencia (E) [47].

El aprendizaje de máquina está tan relacionado a la minería de datos y a la estadística como a la inteligencia artificial y es que no solo trata de aprendizaje, sino de representación de memoria, toma de decisiones e incluso de análisis abstracto de información que está oculta; entonces, si un programa mejora su desempeño en una tarea basado en experiencia pasada, el programa aprendió.

Globalmente se pueden agrupar a los AAM en grupos donde fácilmente pueden encajar según sus funciones y características, distinguiendo 3 principales grupos.

1. El primer grupo de algoritmos son llamados los algoritmos de aprendizaje supervisado, donde el algoritmo intenta producir una salida tomando en cuenta etiquetas de las entradas, un ejemplo común es el agrupamiento selectivo, más conocido por su nombre en inglés Clustering.
2. El aprendizaje no supervisado es el segundo grupo de la clasificación de estos, al no tener etiquetas de entradas el propio algoritmo debe ser capaz de categorizar las mismas, con el fin de producir salidas en base a entradas con patrones encontrados.
3. El tercer grupo de estos algoritmos es el de aprendizaje semi supervisado, combina los dos anteriores, tienen en cuenta entradas etiquetadas y las no etiquetadas para poder clasificar de manera adecuada.

3.2.7.3 APRENDIZAJE PROFUNDO DE MÁQUINA.

El APM puede ser considerado un tipo especial de los AAM. Teóricamente fue desarrollado para que los AAM alcancen un objetivo muy especial, la total independencia (no supervisión) en la toma de decisiones de un proceso o análisis. Pero no siempre los algoritmos de APM son no supervisados.

El APM mundialmente conocido como Deep Learning es entonces solo una rama concreta dentro del Machine Learning (Aprendizaje de Máquina), que agrupa algoritmos conexionistas, los más famosos son las redes neuronales, que permiten la propagación de datos (hacia adelante y recurrentes “acción de que algo vuelva a aparecer”) y la independencia buscada [48].

3.2 HERRAMIENTAS DIGITALES

En la actualidad existen diferentes herramientas para desarrollar las RN, y la actual facilidad de implementación de estas, mediante diferentes librerías como Tensor Flow, Pandas y su módulo Keras en Python™ ofrecen una solución matemática muy superior a la estadística convencional.

3.3.1 ArcGis®

Es una herramienta de geo procesamiento, que facilita el manejo y análisis de información geográfica desarrollado por la compañía ESRI. Sus distintos “toolboxes” y la capacidad de programar funciones en lenguaje Python™ hacen que sea un software robusto para distintas tareas. En la Figura16 se muestra la interfaz gráfica de ArcGis® [49].

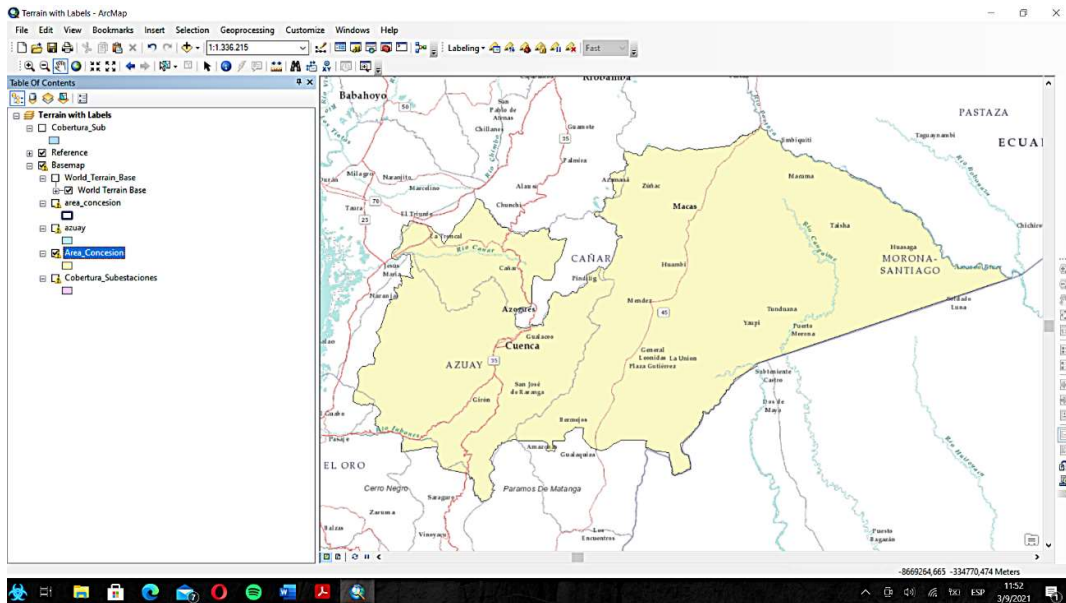


Figura 13. Entorno gráfico de ArcGis®.

El geoprocésamiento proporciona un amplio conjunto de herramientas para realizar tareas GIS que van desde simples zonas de influencia y superposiciones de polígonos a complejos análisis de regresión y clasificación de imágenes. Los tipos de tareas que se van a automatizar pueden ser rutinarias, por ejemplo, alternar un grupo de datos de un formato a otro o las tareas pueden ser complejas, como, por ejemplo, calcular las rutas óptimas a través de la red de transporte, predecir la ruta de un incendio, analizar y buscar patrones en ubicaciones de delitos, predecir qué áreas son propensas a derrumbes o predecir los efectos de inundación de una tormenta [49].

3.3.2 Python™/Pytorch.

Python™ es un lenguaje de programación de código abierto, ya que soporta la programación orientada a objetos, programación imperativa y funcional. Administrado por Python Software Foundation que desde su creación en los años 80 su evolución ha sido significativa hasta la actualidad, ya que es uno de los lenguajes de programación más usado.

Existen distintas librerías enfocadas a distintas tareas, para este análisis se usa la librería Pytorch enfocada en machine learning, redes neuronales. Es una alternativa emergente que está ganando tracción rápidamente gracias a su facilidad de uso, robustez, rapidez gracias a su capacidad nativa para ejecutar en la GPU o tarjeta gráfica, lo que permite acelerar procesos tradicionalmente lentos como el entrenamiento de modelos. Es la librería principal de Facebook para aplicaciones de aprendizaje profundo. Sus elementos fundamentales son los tensores, que se pueden equiparar con vectores de una o varias dimensiones [50].

3.3 CRITERIOS PARA LA EVALUACIÓN DE LA METODOLOGÍA.

La precisión de las redes neuronales y lógica difusa puede ser evaluada por medio de distintos criterios, para el estudio realizado se analizó mediante la aplicación del error cuadrático medio (RMSE), porcentaje de error medio absoluto (MAPE) y el error Kappa, cada criterio corresponde al error existente entre dos conjuntos de datos (real y proyectado) la diferencia entre los valores de test y los proyectados por los algoritmos.

3.4.1 PORCENTAJE DE ERROR MEDIO ABSOLUTO.

Este criterio de evaluación hace referencia al error porcentual, es decir, indica qué tan grandes son los errores de pronóstico en comparación con los valores reales, se calcula hallando el error absoluto para cada uno de los períodos, multiplicado por el inverso del valor real del período correspondiente, y dividido entre N (número de salidas evaluadas) [51] como se indica en la Ecuación 24.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|Real_i - Predicho_i|}{Real_i} \text{ Ecuación 22}$$

3.4.2 ERROR MEDIO CUADRÁTICO (MSE-MEAN SQUARE ERROR)

Este valor se obtiene realizando la sumatoria de cada residuo del período elevado al cuadrado, y dividiendo entre N como indica en la Ecuación 25. Para que el pronóstico sea ideal se espera que el resultado del EMC tienda a 0.

$$RMSE = \frac{1}{N} \sum_{i=1}^N |Real_i - Predicho_i|^2 \text{ Ecuación 23}$$

3.4.3 COEFICIENTE KAPPA.

El coeficiente de Kappa evalúa la similitud entre dos variables, principalmente la comparación entre dos mapas basado en la comparación entre pares de celdas. Por lo tanto, los pequeños desplazamientos entre los mapas se consideran errores y se informa el mismo error incluso si el desplazamiento es de n celdas o una celda [52].

Tabla 4. Niveles de confianza para el valor de k.

Valores de k	Fuerza de la concordancia
< 0.20	Pobre
0.21-0.40	Débil
0.41-0.60	Moderada
0.61-0.80	Buena
0.81-1.00	Muy Bueno

El factor kappa se lo puede definir como en la Ecuación 26:

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \text{ Ecuación 24}$$

Donde Pr (a) la probabilidad de que una celda sea clasificado por los dos observadores o evaluadores en la misma categoría, y Pr (e) es la probabilidad de que los dos observadores concuerden por razones exclusivamente atribuibles al azar [53].

En consecuencia, el valor de k simplemente es la razón entre el exceso de concordancia observado más allá del atribuible al azar y el máximo exceso posible. Si los evaluadores están completamente de acuerdo, entonces k = 1. Si no hay acuerdo entre los calificadores distinto al que cabría esperar por azar $\kappa = 0$. Por lo que la Tabla 3 define los niveles de concordancia del valor k.

3.4.4 FACTOR DE CORRELACIÓN DE PEARSON

El coeficiente de correlación de Pearson es una prueba que mide la relación estadística entre dos variables continuas. Si la asociación entre los elementos no es lineal, entonces el coeficiente no se encuentra representado adecuadamente [54].

El coeficiente de correlación puede tomar un rango de valores de +1 a -1. Un valor de 0 indica que no hay asociación entre las dos variables. Un valor mayor que 0 indica una asociación positiva. Es decir, a medida que aumenta el valor de una variable, también lo hace el valor de la otra. Un valor menor que 0 indica una asociación negativa; es decir, a medida que aumenta el valor de una variable, el valor de la otra disminuye [54], la Ecuación 27 indica como calcular el factor de correlación Pearson.

$$r_{xy} = \frac{\sum z_x z_y}{N} \text{ Ecuación 25}$$

Donde:

N es el número total de muestra.

z_x Desviación estándar de la variable x .

z_y Desviación estándar de la variable y .

3.4.5 R^2

La regresión lineal simple parte de una sola variable predictora, es decir $X=x_1$ donde $X \in \mathbb{R}$ y supone que existe aproximadamente una relación lineal entre X e Y . Esta relación lineal puede ser escrita como en la Ecuación 28 [55]:

$$Y = \beta_0 + \beta_1 X + \epsilon \text{ Ecuación 26}$$

Donde:

ϵ error cometido por el modelo.

β_0 coeficientes de la regresión lineal 0.

β_1 coeficientes de la regresión lineal 1.

3.4.6 Métrica Recall_Score

La métrica Recall Permite calcular la recuperación. Es decir, el retiro o la recuperación es la proporción donde está el número de verdadero positivo y el número de falsos negativos. El retiro es intuitivamente la capacidad del clasificador de encontrar todas las muestras positivas. El mejor valor es 1 y el peor valor es 0 [56]. Cómo se indica en la Ecuación 29:

$$tp / (tp + fn) \text{ Ecuación 27}$$

Donde:

tp verdaderos positivos

fn falsos negativos

3.4.7 Métrica f1_Score

$f1_score$ también conocido como puntuación F equilibrada o medida F. La puntuación F1 se puede interpretar como un promedio ponderado de la precisión y la recuperación, donde una puntuación F1 alcanza su mejor valor en 1 y la peor puntuación en 0. La contribución relativa de precisión y recuperación a la puntuación F1 son iguales. La Ecuación 30 muestra el cálculo para la puntuación F1 es [16]:

$$F1 = 2 * (precision * recall) / (precision + recall) \text{ Ecuación 28}$$

3.4 INTERPOLACIÓN DE PONDERACIÓN DE DISTANCIA INVERSA

La interpolación de ponderación de distancia inversa (IDW) estima valores desconocidos especificando la distancia de búsqueda, los puntos más cercanos, la configuración de potencia y las barreras [57]. La Figura 17 indica gráficamente como se realiza una interpolación (IDW).

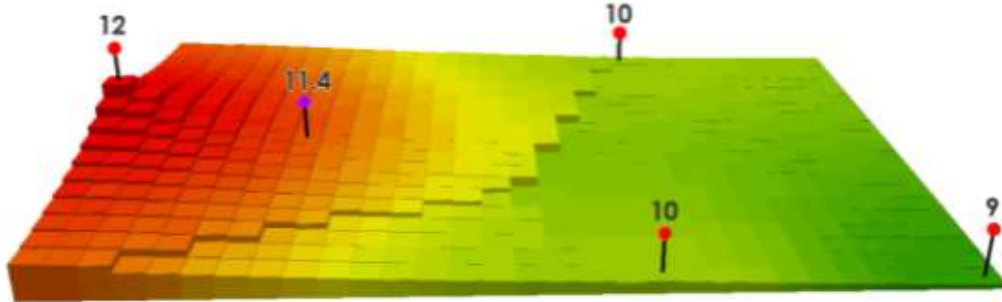


Figura 14. Interpolación ponderada de distancia Inversa. Fuente: [57].

Para realizar la interpolación (IDW), se tiene valores conocidos (valores con marcación roja) y se estima los puntos desconocidos mediante interpolación (valor con marcación morada) como en la Figura 17. La interpolación de ponderación de distancia inversa (IDW) es matemática (determinista) asumiendo que los valores más cercanos están más relacionados que los valores adicionales con su función [57].

La notación sigma simplemente significa que está agregando cualquier número de puntos que se interpolarán. La estimación se la realiza sumando los valores de elevación en cada punto para la distancia.

Un número más pequeño en el denominador (más distancia) tiene menos efecto sobre el valor interpolado. Además, nunca tendrá valores por encima o por debajo de sus valores máximos y mínimos conocidos, se muestra la fórmula matemática en la Ecuación 31 para calcular IDW.

$$\sum p = \frac{\sum_{i=1}^n \left(\frac{z_i}{d_i^p} \right)}{\sum_{i=1}^n \left(\frac{1}{d_i^p} \right)} \text{ Ecuación 29}$$

3.5 MODELOS PARA LA PREDICCIÓN DE FALLAS Y DEFINICIÓN DE LOS MODELOS

Se definen dos modelos de predicción de fallas conocidos como MLP Regresor capaz de trabajar con números continuos y un modelo MLP Classifier que trabajará con números discretos es decir grupos clasificados.

El MLP es una generalización del Perceptron Simple y surgió como consecuencia de las limitaciones de dicha arquitectura en lo referente al problema de la separabilidad no lineal. Diferentes autores han demostrado independientemente que este es un aproximador universal, en el sentido de que cualquier función continua en un espacio R^n puede aproximarse con un MLP, con al menos una capa oculta de neuronas [58].

3.6.1 MLP REGRESSOR.

MLP Regresor entrena iterativamente los modelos neuronales, ya que en cada iteración se calculan las derivadas parciales de la función de pérdida con respecto a los parámetros del modelo para actualizar los parámetros. También se puede agregar un término de regularización a la

función de pérdida que reducen los parámetros del modelo para evitar el sobreajuste. Esta implementación funciona con datos representados como matrices densas y dispersas de valores de punto flotante [59].

3.6.2 MLP CLASSIFIER

MLP Classifier implementa un algoritmo de perceptrón multicapa (MLP) que se entrena mediante la retro propagación. MLP se entrena en dos arreglos: arreglo X, que contiene las muestras de entrenamiento representadas como vectores de características de punto flotante; y matriz Y, que contiene los valores objetivo (etiquetas de clase) para las muestras de entrenamiento. Después del ajuste (entrenamiento), el modelo puede predecir etiquetas para nuevas muestras [60].

3.6 K-MEANS

La clasificación con K-Means para MLP Classifier el algoritmo trabaja iterativamente para asignar a cada "punto" (las filas del conjunto de entrada forman una coordenada) uno de los "K" grupos basado en sus características. Son agrupados en base a la similitud de sus características (las columnas) [61].

IV CAPÍTULO

METODOLOGÍA Y APLICACIÓN DEL MODELO

4.1 INTRODUCCIÓN

Los datos históricos (interrupciones) implican un proceso de registro que contiene información desde el año 2014 hasta el año base 2020. Debido a que juega un papel fundamental en la ingeniería de proyección aplicada, las instituciones encargadas de dicha información son Universidad Politécnica Salesiana (estaciones meteorológicas), Etapa (estaciones meteorológicas), y la Empresa Eléctrica Regional Centro Sur C.A (interrupciones no programadas). La información recolectada permitirá crear datos de interpolación para las variables externas, siendo esta información base para la proyección tendencial realizada hasta el año 2020. La metodología aplicada para Redes Neuronales será la de MLP Regressor y MLP Classifier.

4.2 RECOLECCIÓN DE DATOS

4.2.1 INFORMACIÓN DE INTERRUPCIONES CENTROSUR.

Las interrupciones que se encuentran en la base de datos pertenecientes a la empresa eléctrica regional CENTROSUR, corresponden a las fallas ocurridas dentro del área de concesión de la CENTROSUR pertenecientes a Morona Santiago, Cañar y Azuay, los puntos de fallas corresponden tanto a fallas internas no programadas, internas programadas, externas, vacías, no presenta dato (N/D). Cada una de las interrupciones se clasifica con una causa y subcausa por parte de la empresa y de acuerdo a las causas de interrupciones IEEE 1782. Dentro de la información asociada a cada interrupción se cuenta con información detallada como sus coordenadas, fecha de inicio y fin de la interrupción, duración, así como el alimentador. Esta última información es importante pues permite ubicar con la falla y así asociarla con posibles eventos que la causen.

En la base de datos de interrupciones de la CENTROSUR se filtró únicamente las interrupciones no programadas desde el año 2015 al 2019 ya que en los años 2014 y 2020 no se presentó información.

En el Anexo B se puede observar una muestra de las entradas para la red neuronal de las interrupciones no programadas.

4.2.2 INFORMACIÓN DE ESTACIONES METEOROLÓGICAS

Se realizó una recolección de información sobre las diferentes estaciones ubicadas dentro del área de concesión de la CENTROSUR obteniendo la información dentro de la zona del Azuay con 18 estaciones (Baños, CTS, Cumbe, LLacao, Molleturo, Nulti, San Joaquin, Sayausi, Tixan, Turi, Chaucha, Irquis, Quingeo, Santa Ana, Sinincay, UPS, Cebollar, Ucubamba) cada una de las estaciones brinda información del año 2014 al 2020 en un lapso de tiempo de cada hora.

La información proporciona permite disponer de variables externas y se consideran entre otras entradas:

4.2.1 Presión atmosférica.

Es el peso de la masa de aire por cada unidad de superficie. Por este motivo, la presión suele ser mayor a nivel del mar que en las cumbres de las montañas, aunque no depende únicamente de la altitud [62].

4.2.2 Precipitación.

Se utiliza para referirse a gotas de agua o cristales de hielo que, al ser sometidas a un proceso de condensación y agregación en las nubes, se hacen más pesadas y caen de la atmósfera al suelo [63].

4.2.3 Humedad relativa.

Indica la cantidad de vapor de agua que se encuentra presente en el aire. Depende, en parte, de la temperatura, ya que el aire caliente contiene más humedad que el frío. La humedad se expresa en forma de tanto por ciento (%) de agua en el aire [62].

4.2.4 Radiación solar global.

La radiación solar es la principal fuente de energía (y prácticamente la única) de la atmósfera terrestre. Esta radiación llega en forma de insolación, es decir, rayos de luz y calor de diferentes longitudes de onda que constituyen el espectro visible (rayos luminosos) o que tienen una longitud de onda menor (rayo ultravioleta) o mayor (rayos infrarrojos). Este calor se define como calor atmosférico [62].

4.2.5 Temperatura ambiente.

La temperatura atmosférica (también conocida como temperatura del aire o ambiental) se refiere al grado de calor específico del aire en un lugar y momento determinados, así como a su evolución temporal y espacial en las distintas zonas climáticas. Este elemento meteorológico permite delimitar los tipos de climas, estando en función directa de la mayor o menor cantidad de radiación solar recibida. Es el indicador de la cantidad de energía calorífica acumulada en el aire. Aunque existen otras escalas para otros usos, la temperatura del aire se suele medir en grados centígrados (°C) [62].

4.2.6 Dirección del viento.

La dirección, depende directamente de la distribución de las presiones, pues aquel tiende a soplar desde la región de altas presiones hacia la de presiones más bajas. Se llama dirección del viento el punto del horizonte de donde viene o sopla [64].

4.2.7 Velocidad del viento.

El viento es aire en movimiento. Este desplazamiento es consecuencia de las diferencias de presión y temperatura entre distintas zonas, y se produce en dirección horizontal a lo largo de la superficie terrestre [64].

4.3 ANÁLISIS DE DATOS DE INGRESO

Como información de entrada al algoritmo se escoge tipos de determinantes relacionadas con la duración de las interrupciones. Se trabaja tanto con las variables externas e información de la base de interrupciones de la CENTROSUR tomado de la base como variables coordenadas X, Y, causas, fecha de desconexión y conexión de la falla, alimentador, duración.

La relación de las entradas con la duración se estima mediante un factor de correlación, Se presenta los resultados de la correlación en la Figura 18.

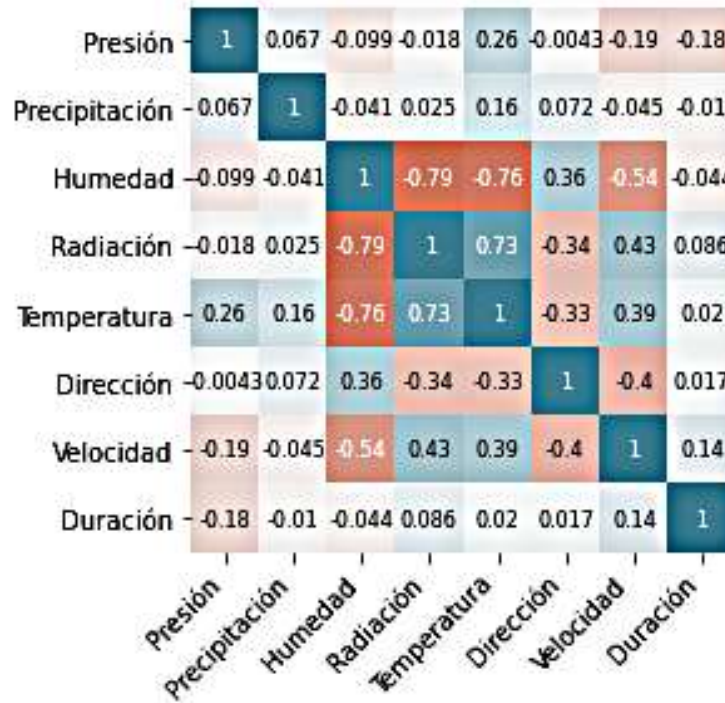


Figura 15. Factor de correlación Pearson.

Como se puede observar el factor de correlación encontrado para cada una de las variables externas en relación con la duración tiene una asociación insuficiente positivamente y negativamente lo ideal sería encontrar variables que tengan mayor asociación llegando a 1.

4.3.1 Correlación entre variables con la duración de la falla.

Mediante la aplicación de la correlación de Pearson entre las variables de estudio con la duración de la falla se tiene la Figura 19.

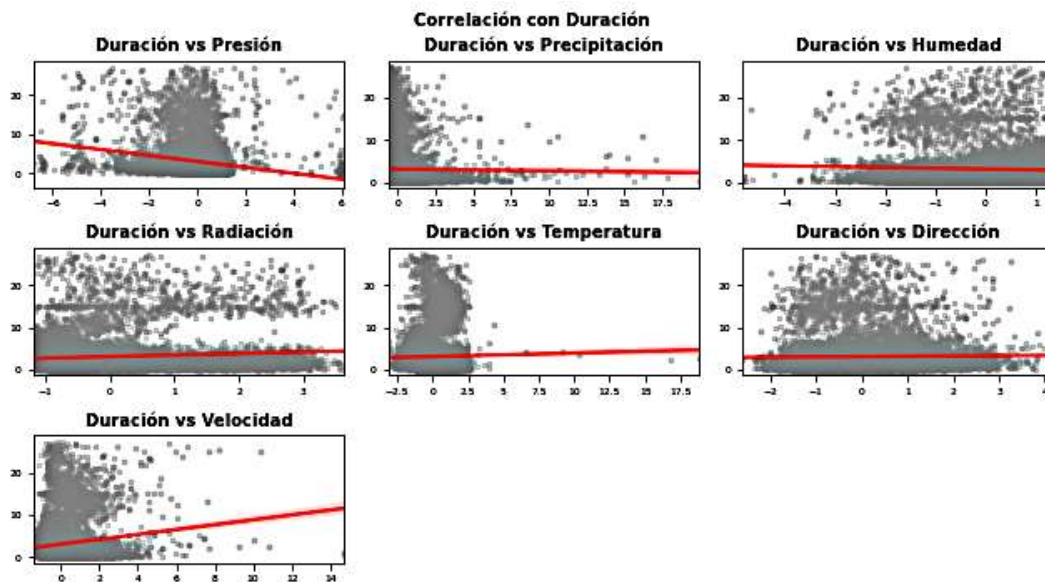


Figura 16. Correlación gráfica de las variables externas con respecto a la duración.

En la Figura 19 se observa la correlación que tiene la duración con las distintas variables externas de esta manera se compara cada uno de los parámetros mostrando gráficamente su relación a lo largo de una línea tendencial roja que representa la correlación de Pearson. Los puntos mostrados representan la base de datos de cada variable externa, teniendo como resultado la relación estadística de las variables continuas que puede ser lineales o no lineales. Para el caso de las 7 variables muestra asociación lineal o una relación directa, positiva en los casos para la radiación, dirección del viento, humedad y presión sin embargo para las variables de precipitación, temperatura y velocidad se relacionan inversamente.

4.4 MAPAS GEOGRÁFICOS

Partiendo de un análisis estadístico de las interrupciones, cuya variable a estratificar es la duración, mediante los procesos realizados el proceso de estratificación de interrupciones nos ayuda a beneficiar la reducción en el tiempo de análisis de los datos y la calidad de la información obtenida además de incrementar la predicción de posibles fallas a futuro.

La información referente a la cantidad de interrupciones y sus variables externas teniendo como resultado para el proceso de estratificación presenta los siguientes resultados, como eliminación de variables sin información, vacíos, entre otros.

4.4.1 UBICACIÓN DE LAS ESTACIONES METEOROLÓGICAS

En el mapa del área de concesión de la CENTROSUR con la información proporcionada de las respectivas ubicaciones de las estaciones meteorológicas, en la Figura 20 se puede observar su ubicación.



Figura 17. Ubicación de las estaciones Meteorológicas.

4.4.2 DEFINICIÓN DE ÁREA

A lo largo del área de concesión de la empresa eléctrica regional CENTROSUR se define el área de las subestaciones 1, 2, 3, 4 5, 7, 8. Estas subestaciones cuentan con la mayor información de

datos de fallas relacionadas con las 18 estaciones meteorológicas definiendo el área que se trabajara como cobertura de subestaciones que se presenta en la Figura 21.

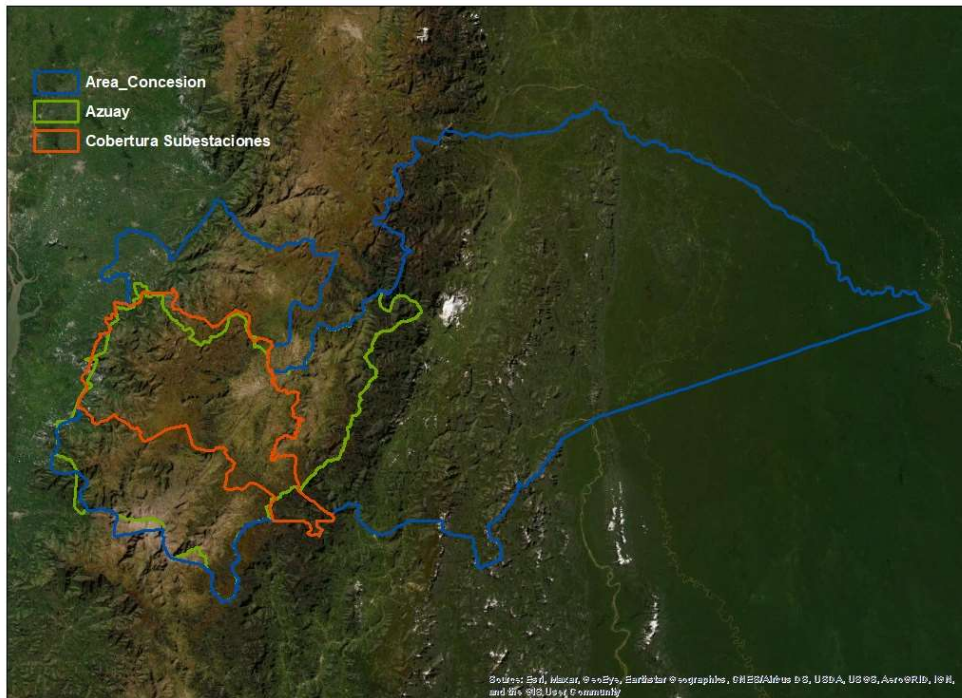


Figura 18. Mapa de Definición de Áreas de CENTROSUR, Área de Concesión, Azuay y Cobertura Subestaciones.

4.5 METODOLOGÍAS APLICADAS, APLICACIÓN DEL MODELO

En la Figura 22 muestra el flujograma para realizar la red neuronal para predecir el tiempo de interrupción de una falla.

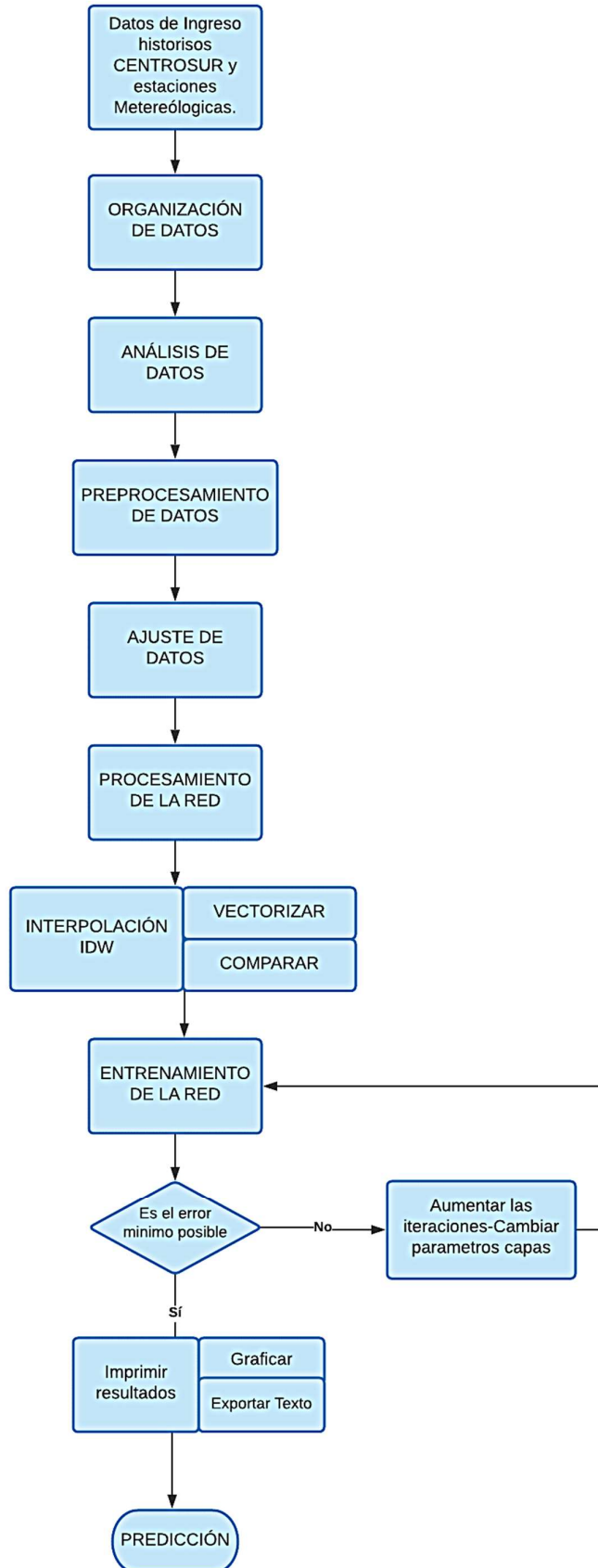


Figura 19. Flujograma del algoritmo de la red neuronal.

4.5.1 ANÁLISIS DE DATOS Y LISTADO DE VARIABLES DE ENTRADA Y SALIDA.

Los datos históricos de las interrupciones no programadas han sido recopilados y proyectados. Se presentan para el estudio dos casos de análisis, para ambos la variable dependiente será la duración difiriendo en las variables independientes como se indica en la Tabla 5.

Tabla 5. Entradas y salidas para diferentes casos de predicción.

	MLP Regressor	MLP Classifier
Entradas	<ul style="list-style-type: none"> • Alimentador • Causas • Coordenada X • Coordenada Y • Presión • Precipitación • Radiación Solar • Humedad • Temperatura • Dirección viento • Velocidad viento • Duración 	<ul style="list-style-type: none"> • Alimentador • Causas • Coordenada X • Coordenada Y • Presión • Precipitación • Radiación Solar • Humedad • Temperatura • Dirección viento • Velocidad viento • Duración
Salidas	Duración	Duración

4.5.2 PREPROCESAMIENTO DE DATOS

El pre procesamiento de datos se puede resumir en las siguientes etapas:

1. Ordenar los datos de todas las estaciones meteorológicas en una hoja sola hoja de cálculo.
2. Crear un nuevo archivo tipo exe. Con él formato de fechas desde el año 2014 al 2020 para poder comparar.
3. Con el software Python™ eliminar datos innecesarios como promedios, mínimos y máximos del archivo de las estaciones meteorológicas.
4. Con el software Python™ comparar los dos archivos e ir cambiando los diferentes tipos de formato y establecer en el formato de dd/mm/aaaa.
5. Reemplaza todos los valores “/”, “ ”, “///”, “-”, “VALID”, “/”, “/// ” por cero.
6. Realizar una comparación tomando los años posteriores o anteriores para completar las celdas vacías. En caso de que no existan datos se realiza un promedio del mes que existe celdas vacías para poder completar y tener la base de datos completa.

4.5.3 AJUSTE DE DATOS

En el área de que corresponde a la cobertura de las subestaciones y con la información de las interrupciones de la CENTROSUR, se establece su ubicación, en función de la localización de las estaciones meteorológicas. En la Figura 23, con la ayuda de las herramientas de ArcMap se establece las fallas en el área de estudio.

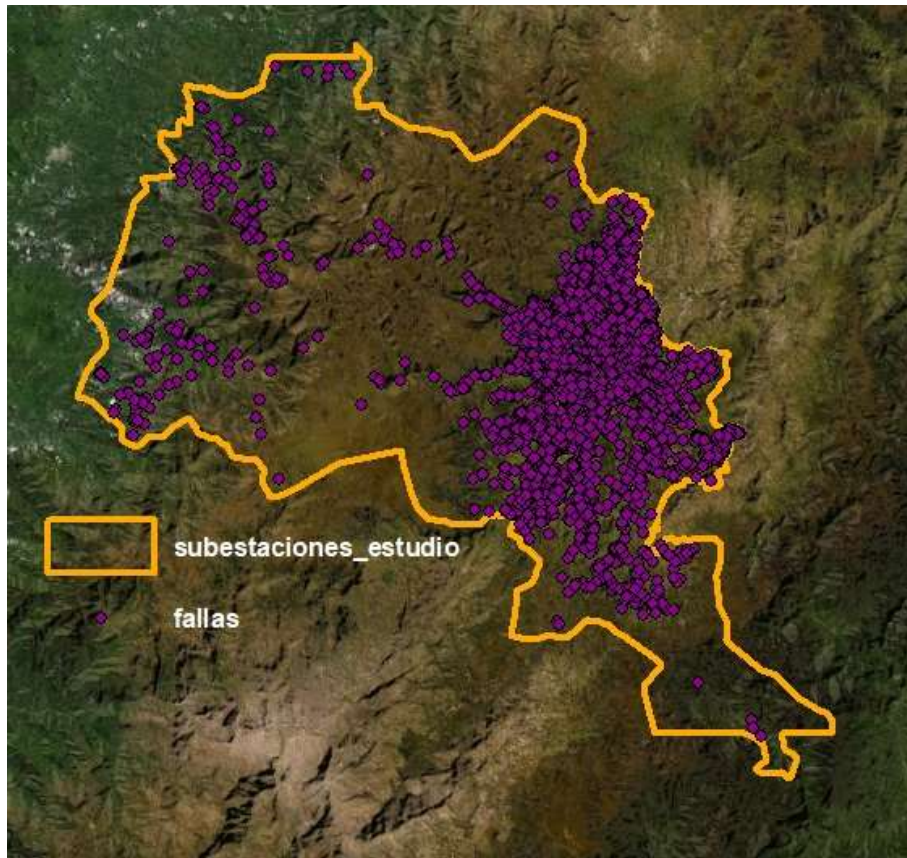


Figura 20. Interrupciones de la CENTROSUR dentro del Área de Cobertura Subestaciones.

Se define un área para poder obtener identificadores de cuadrículas esto debido que se relacionan las interrupciones con las estaciones y con las distancias de identificador de cuadrícula. De esta forma se puede obtener las distancias de cada estación con respecto a cada cuadrícula como se observa en la Figura 24.

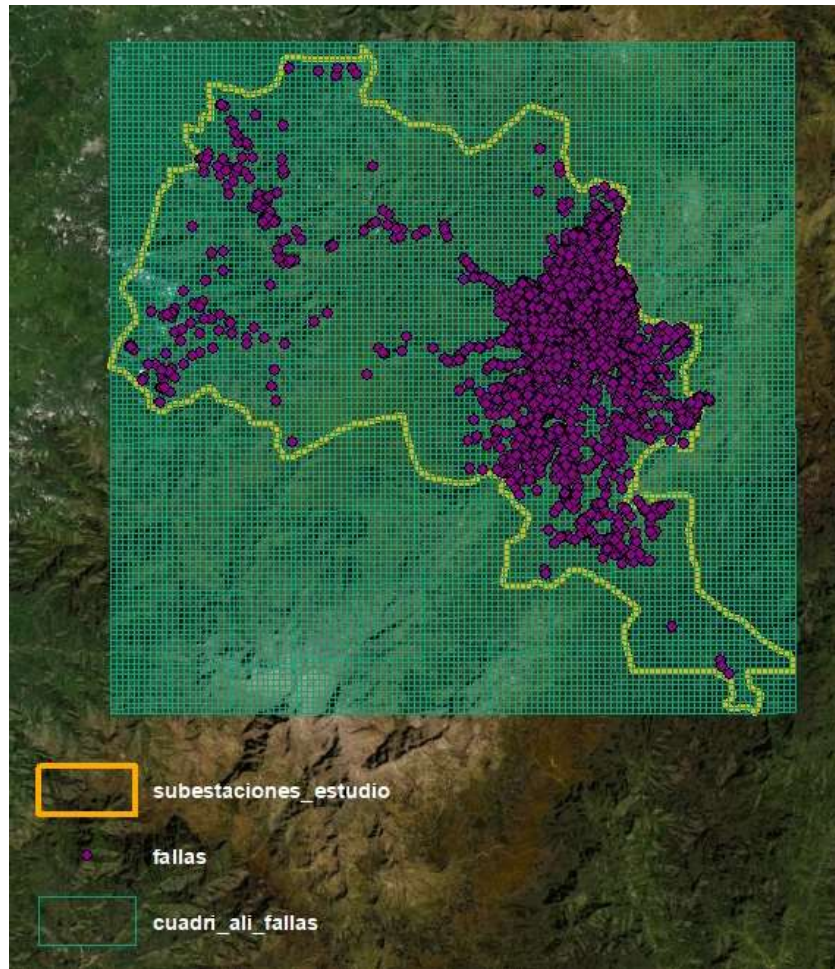


Figura 21. Obtención de distancias mediante las cuadrículas y las estaciones.

Con las herramientas de ArcMap se establecen las distancias de cada punto teniendo a partir de una tabla de atributos la cual se exporta a una hoja de Excel para poder interpolar con la base de datos.

4.5.4 INTERPOLACIÓN DE VARIABLES METEREOLÓGICAS

La interpolación de ponderación de distancia inversa (IDW) permitió distribuir cada una de las variables meteorológicas en el área de estudio. En la Figura 25 se realiza la interpolación (IDW) de la base de datos tomando un dato de las variables externas como es la velocidad del viento en una hora y día específico.

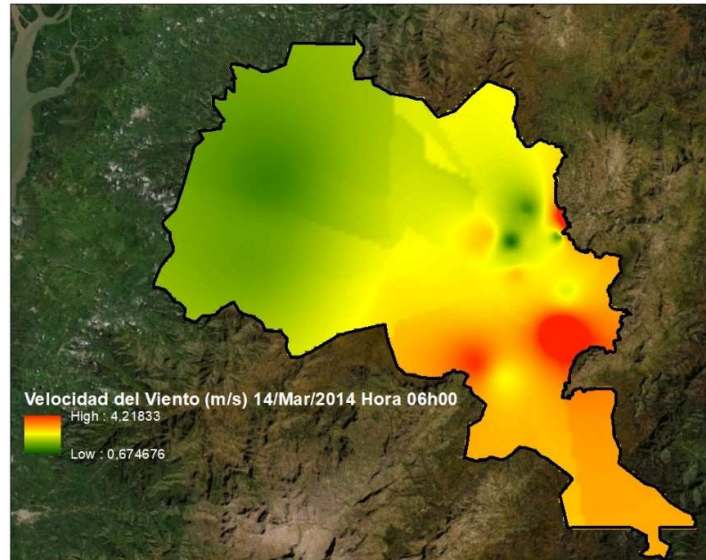


Figura 22. Ponderación de distancia inversa en la velocidad del viento.

Se interpolan los valores para las 7 variables con las distancias a las 18 estaciones.

4.5.5 MODELOS PARA LA PREDICCIÓN DE FALLAS Y DEFINICIÓN DE LOS MODELOS

Para el modelo de predicción se cuenta con 13351 datos de interrupciones con su columna característica para la predicción la variable de duración.

En cada una de las pruebas se modificó el número total de iteraciones, logrando diferentes resultados. Sin embargo, no hay una diferencia marcada por lo cual el número establecido para las iteraciones sea de 1.000.000, de tal manera que a mayor número iteraciones mejor se realiza la predicción, pero el costo computacional entendido como tiempo, crece.

4.5.5.1 MLP REGRESSOR.

Para la red neuronal utilizando MLP Regressor se utilizó específicamente las entradas de las interrupciones y variables externas fijando como objetivo las coordenadas X y Y ya que ayudan con la ubicación de las interrupciones.

En la prueba se consideró epochs al azar igual a 1.000.000, y una configuración: (entrenamiento) = 80% y (pruebas) = 10% y un 10% para validación. Para las diferentes capas ocultas se propuso una estructura (200,150,100,50,25) teniendo una mejor respuesta grafica.

4.5.5.2 MLP CLASSIFIER.

Para la red neuronal utilizando MLP Classifier se utilizó específicamente las entradas de una falla y variables externas fijando como objetivo las coordenadas X y Y ya que permiten ubicar las interrupciones.

En la prueba se consideró epochs al azar igual a 1.000.000, y una configuración: (entrenamiento) = 80% y (pruebas) = 10% y un 10% para validación. Para las diferentes capas ocultas se propuso una estructura (200,150,100,50,25) teniendo una mejor respuesta gráfica.

4.5.6 K-MEANS.

Para determinar el valor de K se realiza una gráfica que permite ubicar el “punto de codo” como se muestra en la Figura 26 donde la duración de la falla fue la variable de entrada.

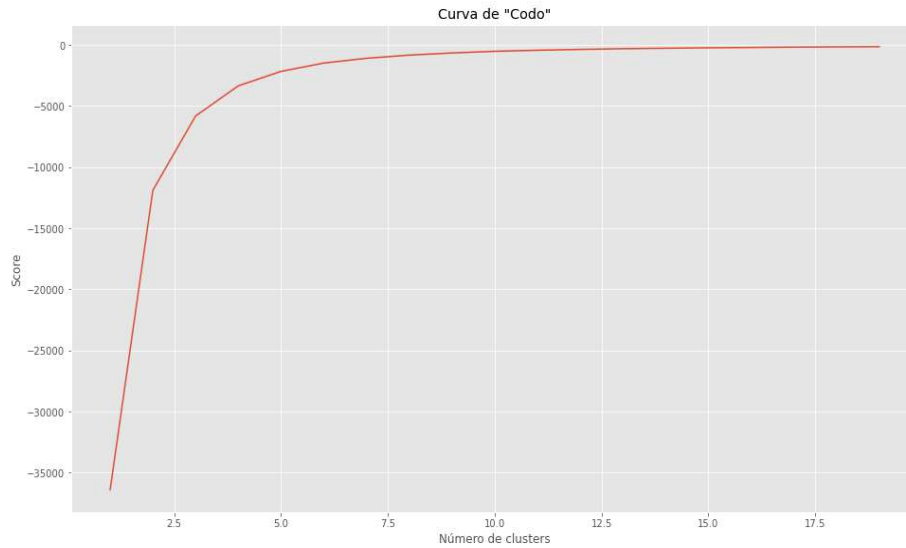


Figura 23. Punto de Codo

Se advierte que la curva es bastante “suave”. Se considera que se estabiliza el valor en 5. Para este método se considera el valor de K como 5 que sería el valor de los grupos en los que se dividirán para el MLP Classifier.

4.6 ENTRENAMIENTO

Para el entrenamiento se ocupará los valores interpolados con IDW que corresponde a la información de las estaciones meteorológicas.

Se realiza un proceso que crea vectores de identificadores para el alimentador, cuadrículas y causas, arregla las fechas y horas de la base de datos final para poder comparar con las de interrupciones.

Considerando la base final de datos que se había dividido en un periodo de 4 horas, para poder minimizar el tiempo computacional, en el momento de realizar la comparación se encontrará con posiciones en la cual se presenta en la Tabla 6, se posicionará en la fecha que engloba a cada posición buscando índice verdadero de la hora.

Tabla 6. Posiciones de las horas para interpolar.

	Posición 1	Posición 2	Posición 3	Posición 4	Posición 5	Posición 6
Hora	1:00h- 4:00h	5:00h- 8:00h	9:00h- 12:00h	13:00h- 16:00h	17:00h- 20:00h	21:00h- 00:00h

4.6.1 FUNCIONES DE ACTIVACIÓN Y OPTIMIZACIÓN

Se evaluó distintas funciones de activación y optimización teniendo como una respuesta para la red, la función de activación Relú y Optimizador Adam en los modelos propuestos como se muestra en la Tabla 7.

Tabla 7. Comparación de las funciones de activación y optimizadores

ACTIVACIÓN	OTIMIZADOR	ITERACIÓN	ERROR RMSE
Identity	Adam	1000000	3,960
Logistic	Adam	1000000	4,040
Tanh	Adam	1000000	4,620
Relu	Adam	1000000	4,620
Identity	Sgd	1000000	3,930
Logistic	Sgd	1000000	4,040
Tanh	Sgd	1000000	4,740
Relu	Sgd	1000000	4,610
Identity	Lbfgs	1000000	3,940
Logistic	Lbfgs	1000000	4,040
Tanh	Lbfgs	1000000	7,370
Relu	Lbfgs	1000000	9,680

En la Figura 27 se muestra los resultados los valores reales-predecidos del test en el modelo MLP Regresor utilizando la función Identity y el optimizador Stochastic Gradient Descent (SGD). Con estas técnicas sé que presento el menor valor en el error RMSE de 3.93.

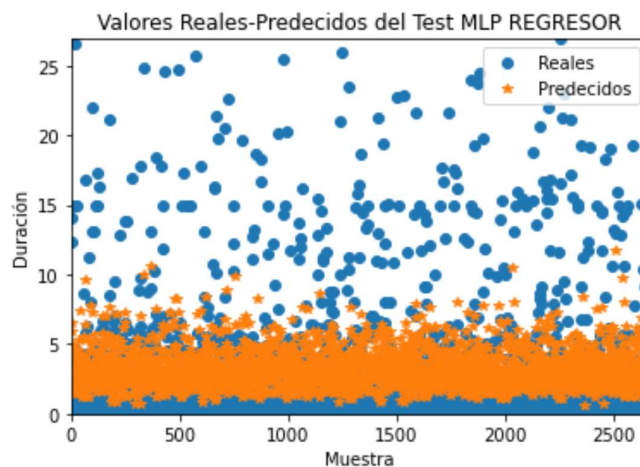


Figura 24. Resultado del entrenamiento con la función de activación Identity y optimizador Sgd.

En la Figura 28 se muestra los resultados los valores reales-predecidos del test en el modelo MLP Regresor utilizando la función Relú y el optimizador Adam, no presenta el menor valor en el error RMSE, pero se representa mejor gráficamente.

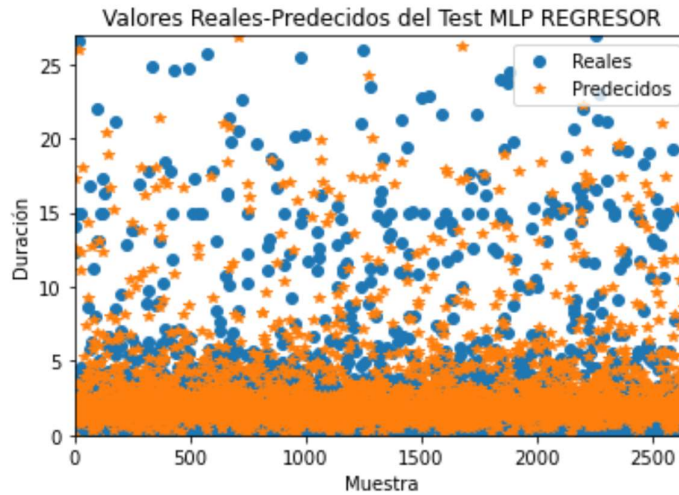


Figura 25. Resultado del entrenamiento con la función de activación Relú y optimizador Adam.

4.7 RESULTADOS

Dentro de los resultados de los diferentes modelos presentados y con la finalidad de diseñar un modelo confiable y aplicable se prevé mejorar la confiabilidad de la predicción y el tiempo que toma realizar el proceso.

4.7.1 MLP REGRESSOR

En la Figura 29 muestra los resultados del entrenamiento del modelo MLP Regressor. Se destaca que existe un buen entrenamiento en el modelo ya que predice en su totalidad todos los valores y como resultado se observa el índice de error es mínimo.

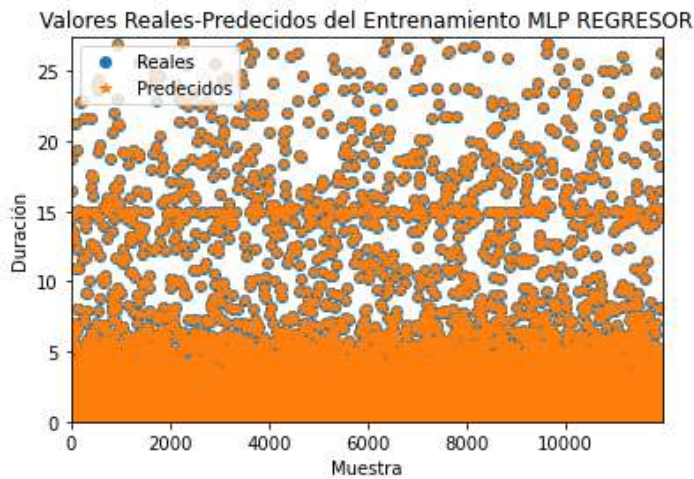


Figura 26. Resultado del Entrenamiento con MLP Regressor.

La Tabla 8 muestra las diferentes métricas aplicas para el modelo MLP Regressor mostrando los índices de errores en el entrenamiento del modelo.

Tabla 8. Resultados de los errores con la red MLP Regressor en el entrenamiento.

MLP REGRESSOR		
Entrenamiento	RMSE	1,4205

	R^2	0,8899
	Error absoluto Medio	0,9372
	Varianza	0,8900

En la Figura 30 muestra los resultados del test del modelo MLP Regresor, indica que ya no existe una buena predicción de los valores y como resultado se observa el índice de error en la Tabla 8.

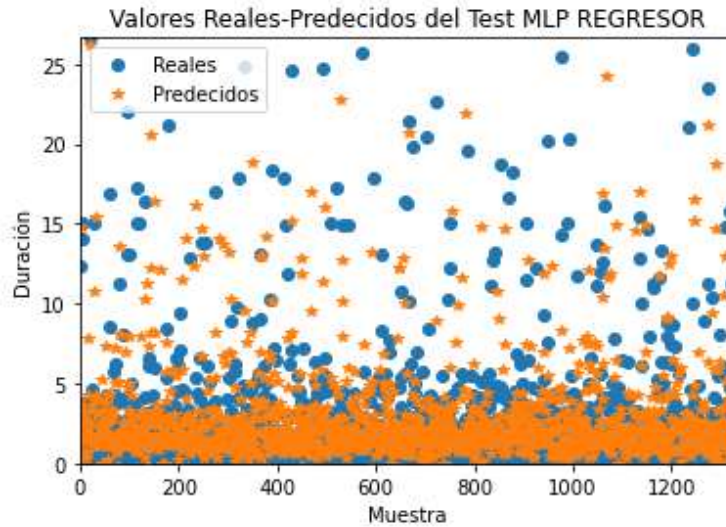


Figura 27. Resultado del Test con MLP Regresor.

La Tabla 9 muestra las diferentes métricas aplicas para el modelo MLP Regresor mostrando los índices de errores en el test del modelo.

Tabla 9. Resultados de los errores con la red MLP Regresor en el test.

Test	RMSE	4.6144
	R^2	0,3870
	Error absoluto Medio	2,5518
	Varianza	0,3870

4.7.2 MLP CLASIFIER

Para el agrupamiento de datos se han tomado dos criterios:

- Dando rangos de acuerdo a la información de la duración en 6 grupos basada en información de expertos (CENTROSUR).
- Mediante el método K-Means.

En la Tabla 10 se establece los tiempos de duración para ambos criterios en diferentes grupos como se muestra:

Tabla 10. División de grupos de interrupción de acuerdo a los criterios establecidos a la duración.

Criterio	Experto	K-Means
GRUPOS	Duración	
G1	De 0 a 5 min	De 0 a 9 min

G2	De 5 a 30 min	De 9 a 26 min
G3	De 30 a 2h	De 26 a 47 min
G4	De 2h a 5h	De 47 a 1.10h
G5	De 5 a 10h	Mayore a 1.10h
G6	Mayor a 10h	

De acuerdo a la clasificación de los grupos por horas se registra en cada grupo un numero de datos como se muestra a continuación en la Tabla 11 para los dos tipos de modelos.

Tabla 11. Tamaño de muestra para los criterios establecidos.

Criterio	Experto	K-Means
GRUPOS	Tamaño de muestra	
G1	442	688
G2	1796	1458
G3	5202	1336
G4	3942	1232
G5	940	8636
G6	1028	

En la Figura 31 se muestra una tabla de distribución del tamaño de muestras para ambos criterios.

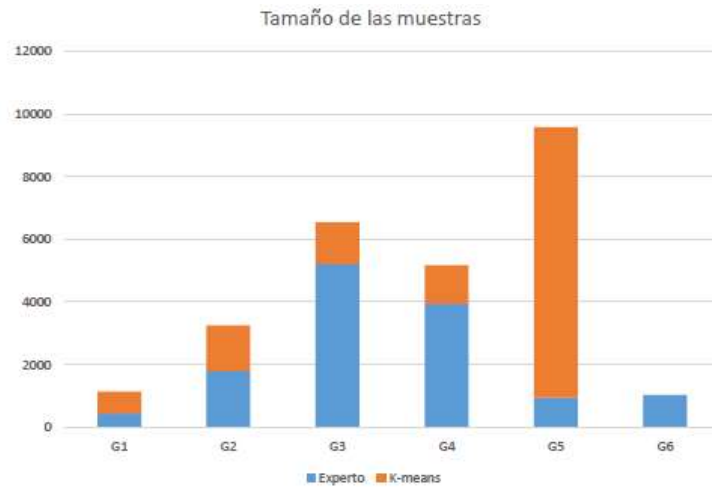


Figura 28. Distribución del tamaño de muestras para criterio de K-Means y de expertos de la CENTROSUR.

La Tabla 12 muestra las diferentes métricas aplicadas para el modelo MLP Classifier utilizando el criterio de agrupamiento de un Experto. Se observan los índices de errores en el entrenamiento.

Tabla 12. Resultados de los errores con la red MLP Classifier en el entrenamiento según el criterio de un experto.

Grupo	Precisión	Recall	F1-Score	Support
G1	0,91	0,95	0,93	405
G2	0,91	0,95	0,93	1622
G3	0,98	0,93	0,95	4688

G4	0,96	0,94	0,95	3526
G5	0,81	0,97	0,88	845
G6	0,92	0,96	0,94	929
Accuracy			0,94	12015
Macro avg	0,91	0,95	0,93	12015
Weighted avg	0,944	0,94	0,94	12015

La Tabla 13 muestra las diferentes métricas aplicas para el modelo MLP Classifier y en la Tabla 14 se muestra el modelo MLP Classifier utilizando K-Means. Así mismo se observa los índices de errores en el entrenamiento y en la Tabla 15 el test del modelo.

Tabla 13. Resultados de los errores con la red MLP Classifier en el entrenamiento.

Grupo	Precisión	Recall	F1-Score	Support
G1	0,20	0,24	0,22	36
G2	0,34	0,45	0,39	174
G3	0,44	0,41	0,42	514
G4	0,42	0,35	0,38	416
G5	0,15	0,18	0,16	95
G6	0,27	0,34	0,30	99
Accuracy			0,37	1335
Macro avg	0,30	0,33	0,31	1335
Weighted avg	0,38	0,37	0,37	1335

Tabla 14. Resultados de los errores con la red MLP Classifier en el entrenamiento según el criterio K-Means.

Grupo	Precisión	Recall	F1-Score	Support
G1	0,89	0,96	0,92	579
G2	0,98	0,89	0,93	1200
G3	0,95	0,86	0,90	1243
G4	0,95	0,92	0,93	1240
G5	0,96	0,99	0,98	7753
Accuracy			0,96	12015
Macro avg	0,95	0,92	0,93	12015
Weighted avg	0,96	0,96	0,96	12015

Tabla 15. Resultados de los errores con la red MLP Classifier en el test según el K-Means.

Grupo	Precisión	Recall	F1-Score	Support
G1	0,18	0,26	0,21	53
G2	0,36	0,32	0,34	126

G3	0,17	0,12	0,14	150
G4	0,16	0,13	0,14	123
G5	0,75	0,80	0,77	883
Accuracy			0,59	1335
Macro avg	0,32	0,33	0,32	1335
Weighted avg	0,57	0,59	0,58	1335

En las Figura 32 se presentan los resultados del test de una muestra de la base de datos del modelo MLP Classifier por el criterio que brindo mejores resultados K-Means. De esta manera se establece que está prediciendo correctamente en la mayoría de datos.

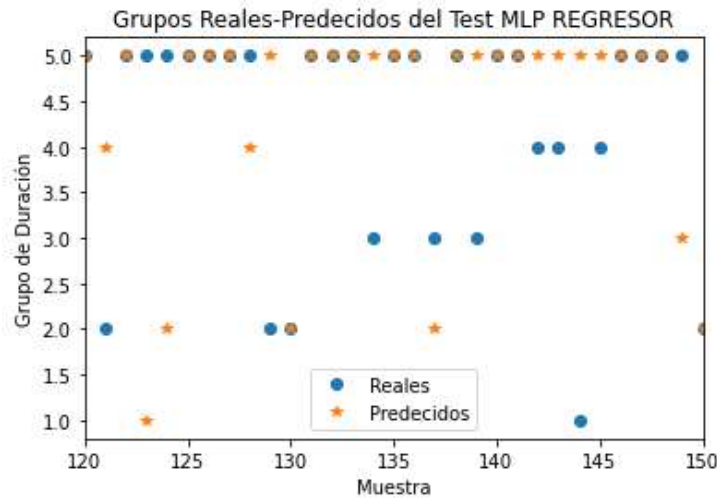


Figura 29. Resultado del Test con MLP Classifier con K-Means.

Dentro de los resultados de los diferentes modelos presentados y con la finalidad de diseñar un modelo confiable y aplicable se prueba la confiabilidad de la predicción y el tiempo que toma realizar el proceso.

Los modelos se probaron con un procesador Intel Core i5, con la finalidad de probar de que se puede mejorar el tiempo de procesamiento en una PC con mejor procesador. Para todos los modelos se tomó un máximo de 1.000.000 iteraciones.

Se evalúa el modelo con una fila tomada desde la base de datos final que se presenta con información presentada en la Tabla 16.

Tabla 16. Información de interrupción de Cortocircuito.

Entradas	Información
Alimentador	521
Fila	4
Causa	CORTOCIRCUITO
Coordenada_X	965.908,70
Coordenada_Y	720.538,19
Presión	715,00
Precipitación	0,31
Humedad	97,25

Radiación	9,99
Temperatura	8,59
Dirección	101,22
Velocidad	3,68
Duración	2,26

Analizando el modelo MLP Regressor como se muestra se concluye que la diferencia no es muy amplia tomando un valor de 2.04 con respecto al valor real de 2.26, El resultado en el modelo MLP Classifier K-Means al igual que el modelo MLP Regressor la diferencia no es muy amplia sin embargo como son modelos de clasificación entra al grupo de un tiempo predicho de 26 a 48 min.

En el caso MLP Regresor se tiene un valor de 4,6144 de margen de error en el test y en el caso MLP Classifier con K-Means presenta un valor de 0.33 de margen de error sin embargo en este modelo presenta una mejor distribución de los datos.

V CAPÍTULO

DISCUSIÓN

5.1 DISCUSIÓN.

Conociendo que los sistemas de distribución constantemente están bajo la amenaza de cortocircuitos y se provee mejorar la calidad y confiabilidad de los sistemas de distribución, las empresas de distribución de energía pretenden despejar rápidamente las fallas de modo que la restauración sea lo más rápida posible.

Para ello se ha planteado diferentes métodos de localización de fallas, así como métodos tradicionales y enfoques de aprendizaje automático conociendo que la mejor opción serían los métodos enfocados al aprendizaje automático sin embargo son difíciles de modelar.

Para este estudio basándonos en varios autores se ha tomado la opción de ir por los métodos de aprendizaje automático, llevando los datos a introducirlos en una red neuronal artificial para la ubicación de la falla como menciona el autor en [7].

El modelo presentado de un MLP presento buenos resultados a comparación del modelo del DNN como indica el autor en [11] teniendo un mejor resultado en el margen de error.

Se utilizó de igual manera regresión lineal considerando todas las causas de las interrupciones, para el análisis también se consideró las variables externas de las estaciones meteorológicas incluyendo fallas temporales y permanentes al igual que indica el autor en [9], esto gracias a la información recopilada de un conjunto extenso de datos históricos.

Según los resultados del proyecto se encontró un número de variables externas asociadas con la duración realizando una correlación de análisis para la mejora de la gestión de interrupciones en las redes con datos meteorológicos como indica el autor en [14], este análisis desempeñan un rol importante y significativo para el proyecto, no solo por la implementación del método predictivo sino también por la aportación.

Además, es de suma importancia utilizar un amplio conjunto de datos históricos con aproximación de 6 años para poder tomar 5 años de entrenamiento y el sexto año para validación como se realizó y menciona el autor en [11]. Se proporcionó un procedimiento detallado, que incluye varios pasos para tratar de manera efectiva la preparación de datos requerida, problemas prácticos asociados con el conjunto de datos de fallas del mundo real y la implementación de MLP [15].

Se presenta en el proyecto una mejora sobre un sistema de predicción de fallas, basado en Logistic Regression y la aplicación de técnicas de predicción de eventos raros, capaz de trabajar con datasets dispersos y de alta dimensión como indica en [16].

Sin embargo, un problema muy común para este proyecto es la escasez de datos que para este tipo de modelos son de vital importancia, la mayor parte de la matriz del conjunto de datos está ausente de eventos, esta situación es complicada para la mayoría de los métodos que se pueda emplear porque al conjunto de datos se tienden a sobre ajustarse o no estimar correctamente la variable objetivo como indica el autor en [16].

La aplicación del modelo con la información de CENTROSUR brindo una mejora considerable con los modelos aplicados en la literatura, esto es debido a: calibración correcta del modelo, series de tiempo más extensas y la aplicación de diferentes modelos como MLP regresor y clasificador. Asegurando un alto grado de confiabilidad para garantizar al cliente una correcta información referente al tiempo de duración de la falla.

VI CAPÍTULO

CONCLUSIONES Y RECOMENDACIONES

6.1 CONCLUSIONES.

En el trabajo de titulación se desarrolló un modelo para predecir el tiempo de duración de las fallas para aquellas subestaciones pertenecientes a la CENTROSUR y que están ubicadas en la ciudad de Cuenca.

Las estaciones meteorológicas con las interrupciones no programadas al incluirse al sistema eléctrico de CENTROSUR juegan un papel importante en la predicción, pues estas ayudan de manera considerable a predecir el tiempo de falla a más de sus variables en las subestaciones involucradas. Debido a estas variables externas en el servicio eléctrico se ve afectado, ya que existen fallas continuas en diferentes años.

Se logró determinar que los modelos de aprendizaje de máquina brindan un buen resultado, pero no lo esperado para el trabajo con predicciones de tiempo. Incluso los valores pronosticados pueden ser comparados con la base de la empresa CENTROSUR. Cabe mencionar que la estimación con APM tuvo como dato de entrada puntos importantes de la base de la Empresa y variables externas de estaciones meteorológicas, demostrando que las metodologías son buenas, pero por sí solas no suficientes.

El criterio de evaluación MAPE implicó ser dominante, debido a que se promedió datos inexistentes de las variables externas, esto se debe a que algunas estaciones no presentaron datos históricos en diferentes años, y falló al haber creado una red con datos promediados.

En la base de datos al considerar valores máximos se utilizó el criterio estadístico de desviación estándar que agrupa los datos debajo del valor hallado. Para este análisis la red aplicó la primera desviación estándar ya que aplicando una segunda o tercera desviación los resultados no favorecieron a la red.

Dentro de los diferentes problemas existentes de predicción relacionados a eventos en el tiempo, los modelos de aprendizaje profundo de máquina son muy positivos y efectivos. Hoy en día existen herramientas que facilitan la construcción de las redes neuronales, como diferentes librerías que se programan a través de plataformas de código abierto para el aprendizaje automático. Estas herramientas crean en pocas líneas de código un ecosistema integral y flexible que permiten innovar, construir cualquier arquitectura compleja de red neuronal de forma sencilla, rápida y precisa.

El desarrollo del modelo computacional para la predicción de fallas usando machine learning con interpolación inversa ponderada en las subestaciones de Cuenca con los diferentes puntos de estaciones meteorológicas resulta de interés teórico y práctico debido a:

- Que las variables de tiempo pueden ser de gran utilidad en las interrupciones debido a que se relacionan a todos los factores en los sistemas de distribución.
- Que permiten comprobar que no es óptimo realizar una exuberante cantidad de iteraciones para conseguir resultados positivos en un proceso.
- Que el aprendizaje resulta mejor si se realiza una clasificación de valores es decir, agrupando valores máximos y mínimos para mejores resultados en la red neuronal.

Se presentó un mejor resultado aplicando una clasificación (MLP Classifier, K-Means) y se constató que una función Relú aparte de ser muy popular estos días, es de beneficio en la salida de una RMLP.

Los resultados de esta investigación muestran que es posible desarrollar estos sofisticados modelos prácticamente en cualquier empresa donde se cuente con información recolectada. Esto puede contribuir al desarrollo de Big Data, tan importante en esta década, sin contar que hay muchas herramientas de Software libre que permiten el desarrollo de diferentes aplicaciones.

6.2 RECOMENDACIONES

Los resultados en los distintos análisis son semejantes en las predicciones, aún se puede realizar mejoras en trabajos futuros, como son: combinación de los métodos para ayudar el proceso de aprendizaje y pronóstico.

Para el entrenamiento de la red se requiere mayor cantidad de datos históricos y en lo posible adicionar nuevas variables de ingreso que tengan un aceptable factor de correlación con la duración tales como factores climáticos, demográficos, geográficos.

Para mejorar la confiabilidad de la predicción es recomendable usar datos de fuentes u organismos oficiales, logrando de esta manera reducir el error evaluado y que se encuentre dentro de los parámetros adecuados. Así mismo, se puede plantear distintos escenarios en cuanto a variables independientes para tener una visión global en multiescenarios.

En el proceso de proyección es recomendable no considerar datos que son extremadamente grandes o pequeños debido a que generan datos atípicos en los históricos afectando directamente a la fase de entrenamiento. Por lo que, se recomienda excluir esta información para luego adicionarla en los resultados.

El APM no ha sido explorado por la industria eléctrica en problemas similares, por lo cual el uso de APM para resolver el problema de predicción de fallas resulta de interés de las compañías de distribución eléctrica locales e internacionales para aplicaciones pioneras.

La CENTROSUR presenta un registro de fallas, algunas repetitivas, todos los meses del año y actualmente no cuenta con ninguna herramienta tecnológica que sea capaz de predecir la ocurrencia de fallas. Esta herramienta podría aportar mucho a la empresa de manera económica.

Los sistemas para la recopilación de información como el SCADA, toman mayor protagonismo a la hora de analizar grandes cantidades de datos. Por esto, la utilidad de este trabajo resulta relevante si esos datos pueden servir para la creación de algoritmos de predicción.

Finalmente, se recomienda que la Empresa haga énfasis en el registro de datos de las interrupciones después de que sufre una falla, debido a la gran cantidad de datos restantes en la base de datos.

BIBLIOGRAFÍA

- [1] S. Das, N. Karnik, and S. Santoso, "Distribution Fault-Locating Algorithms Using Current Only," *IEEE Trans. Power Deliv.*, vol. 27, no. 3, pp. 1144–1153, Jul. 2012, doi: 10.1109/TPWRD.2012.2191422.
- [2] R. A. F. Pereira, L. G. W. da Silva, M. Kezunovic, and J. R. S. Mantovani, "Improved Fault Location on Distribution Feeders Based on Matching During-Fault Voltage Sags," *IEEE Trans. Power Deliv.*, vol. 24, no. 2, pp. 852–862, Apr. 2009, doi: 10.1109/TPWRD.2009.2014480.
- [3] Y. Jiang, C.-C. Liu, M. Diedesch, E. Lee, and A. K. Srivastava, "Outage Management of Distribution Systems Incorporating Information From Smart Meters," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4144–4154, Sep. 2016, doi: 10.1109/TPWRS.2015.2503341.
- [4] "Fuzzy spatial load forecasting."
- [5] D. F. Rodríguez Rodríguez and C. C. Guevara Serrano, "Metodología de localización de fallas basada en ondas viajeras con señales de tensión - hdl: 11349/13856."
- [6] D. Thukaram, H. P. Khincha, and H. P. Vijaynarasimha, "Artificial Neural Network and Support Vector Machine Approach for Locating Faults in Radial Distribution Systems," *IEEE Trans. Power Deliv.*, vol. 20, no. 2, pp. 710–721, Apr. 2005, doi: 10.1109/TPWRD.2005.844307.
- [7] Y. Aslan and Y. E. Yağan, "Artificial neural-network-based fault location for power distribution lines using the frequency spectra of fault data," *Electr. Eng.*, vol. 99, no. 1, pp. 301–311, Mar. 2017, doi: 10.1007/s00202-016-0428-8.
- [8] M. Dokic, Tatjana;Chen, Po Chen;Kezunovic, "Risk Analysis for Assessment of Vegetation Impact on Outages in Electric Power Systems," *CIGRE US Natl. Comm. 2016 Grid Futur. Symp.*, no. May, pp. 1–23, 2013.
- [9] R. D. M. KEZUNOVIC, T. DOKIC , Z. OBRADOVIC, M. PAVLOVSKI, "Big Data Analytics for Predictive Lightning Outage Management Using Spatially Aware Logistic Regression Model."
- [10] A. Moradkhani, M. R. Haghifam, and M. Mohammadzadeh, "Bayesian estimation of overhead lines failure rate in electrical distribution systems," *Int. J. Electr. Power Energy Syst.*, vol. 56, pp. 220–227, 2014, doi: 10.1016/j.ijepes.2013.11.022.
- [11] A. Arif and Z. Wang, "Distribution network outage data analysis and repair time prediction using deep learning," *2018 Int. Conf. Probabilistic Methods Appl. to Power Syst. PMAPS 2018 - Proc.*, pp. 1–6, 2018, doi: 10.1109/PMAPS.2018.8440354.
- [12] M. Doostan, R. Sohrabi, and B. Chowdhury, "A data-driven approach for predicting vegetation-related outages in power distribution systems," *Int. Trans. Electr. Energy Syst.*, vol. 30, no. 1, Jan. 2020, doi: 10.1002/2050-7038.12154.
- [13] J. B. Leite, J. R. S. Mantovani, T. Dokic, Q. Yan, P. C. Chen, and M. Kezunovic, "Failure probability metric by machine learning for online risk assessment in distribution networks," *2017 IEEE PES Innov. Smart Grid Technol. Conf. - Lat. Am. ISGT Lat. Am. 2017*, vol. 2017-Janua, pp. 1–6, 2017, doi: 10.1109/ISGT-LA.2017.8126683.
- [14] P. Chen, "THE USE OF BIG DATA FOR OUTAGE MANAGEMENT IN DISTRIBUTION SYSTEMS II . BACKGROUND : DATA SOURCES USED CIRED Workshop - Rome , 11-12 June 2014 Paper 0406 A . The Use of Different FL Techniques B . The Use of Customer Calls and Smart," no. 0406, pp. 11–12, 2014.
- [15] M. Doostan and B. H. Chowdhury, "Power distribution system fault cause analysis by

- using association rule mining,” *Electr. Power Syst. Res.*, vol. 152, pp. 140–147, 2017, doi: 10.1016/j.epsr.2017.07.005.
- [16] J. M. Navarro, A. P. G. Hugo, and J. C. Dueñas, “Classification in sparse, high dimensional environments applied to distributed systems failure prediction,” *Lect. Notes Artif. Intell. (Subseries Lect. Notes Comput. Sci.)*, vol. 9120, pp. 714–726, 2015, doi: 10.1007/978-3-319-19369-4_63.
- [17] Empresa Electrica Regional CENTRO SUR C.A, “PLAN ESTRATEGICO 2017-2021,” Cuenca. Accessed: Jun. 14, 2021. [Online]. Available: <https://www.centrosur.gob.ec/wp-content/uploads/2018/10/PLAN-ESTRATEGICO-INSTITUCIONAL-2017-2021-APROBADO.pdf>.
- [18] Instituto Ecuatoriano de Normalización, “NORMA TECNICA ECUATORIANA ‘NTE INEN-IEC 61000-2-2,’” *IEC 2020*, Jan. 2014. https://www.normalizacion.gob.ec/buzon/normas/NTE_INEN_IEC_61000_2_2.pdf (accessed Sep. 30, 2020).
- [19] “CENELEC - Desarrollo de Normas - Relación de Órganos Técnicos -.” https://www.cenelec.eu/dyn/www/f?p=104:110:1031043823592201:::FSP_ORG_ID,FS_P_PROJECT,FSP_LANG_ID:1258595,51993,25 (accessed Sep. 30, 2020).
- [20] CENTROSUR, “Interrupciones anteriores.” <https://www.centrosur.gob.ec/interrupciones-anteriores/> (accessed Oct. 19, 2020).
- [21] D. T. Radmer, P. A. Kuntz, R. D. Christie, S. S. Venkata, and R. H. Fletcher, “Predicting vegetation-related failure rates for overhead distribution feeders,” *IEEE Trans. Power Deliv.*, vol. 17, no. 4, pp. 1170–1175, Oct. 2002, doi: 10.1109/TPWRD.2002.804006.
- [22] J. F. Segarra and P. S. Andrade, ““MODELO PREDICTIVO DE FALLAS EN ALIMENTADORES PRIMARIOS DE CONCESIÓN DE LA EMPRESA ELÉCTRICA REGIONAL CENTRO SUR USANDO APRENDIZAJE PROFUNDO DE MÁQUINA,”” Universidad de Cuenca, 2017.
- [23] Aristóteles and J. D. Juárez Cervantes, *Sistemas de Distribución de Energía Eléctrica*, vol. 52, no. 1. 2002.
- [24] G. Hernández, “Electricista y Transitorios,” Jun. 25, 2018. <https://blogspanol.se.com/uncategorized/2018/06/25/electricista-listo-para-las-lluvias-y-los-transitorios/> (accessed Jun. 15, 2021).
- [25] P. J. Seymour, “Siete tipos de problemas en el suministro eléctrico,” *Am. Power Convers.*, pp. 6–20, 2005.
- [26] H. T. Seymour Joseph, “los siete tipos de problemas en el suministro eléctrico.” Accessed: Jul. 01, 2021. [Online]. Available: https://download.schneider-electric.com/files?p_Doc_Ref=SPD_VAVR-5WKLPK_LS.
- [27] J. Aníbal, E. Gualli, W. Patricio, and O. Trujillo, “METODOLOGÍA PARA LA EVALUACIÓN DE LA CALIDAD DE SERVICIO TÉCNICO DE LA EMPRESA ELÉCTRICA COTOPAXI,” ESCUELA POLITECNICA NACIONAL, COTOPAXI, 2001.
- [28] F. Perez and E. Vaca, “DESARROLLO DE UNA HERRAMIENTA COMPUTACIONAL PARA LA SIMULACION DE METODOS DE POLARIZACION DE LA PROTECCION DE SOBRECORRIENTE DIRECCIONAL DE LINEAS DE TRANSMISION,” 2019.
- [29] B. B. Armengol, *Protección De Sistemas De Potencia*. 2018.

- [30] C. Alberto, A. Avendaño, H. Felipe, and I. Olaya, “Sobretensiones transitorias en sistemas eléctricos,” *Tecnura*, vol. 4, no. 7, pp. 43–48, Jul. 2000, doi: 10.14483/22487638.6105.
- [31] O. P. Ardila, M. Iván, U. Gamboa, W. Leonardo, and R. Galvis, “Equipo Analizador de Calidad de Energía Eléctrica implementado en Matlab y Paquete de Adquisición de Datos .”
- [32] *Norma de calidad de potencia eléctrica aplicables a los servicios de Distribución de Energía Eléctrica*. . CREG.
- [33] “Perturbaciones en la onda de tensión: Huecos [sag] y Sobretensiones [swell] - PDF Free Download.” <https://docplayer.es/19962992-Perturbaciones-en-la-onda-de-tension-huecos-sag-y-sobretensiones-swell.html> (accessed Feb. 01, 2022).
- [34] H. Pila and M. Zambrano, “Estudio del Flicker en una Instalación Eléctrica,” Universidad Politecnica Salesiana, Guayaquil, 2010.
- [35] J. Seymour and T. Horsley, “Los siete tipos de problemas en el suministro eléctrico.” Accessed: Apr. 19, 2021. [Online]. Available: https://download.schneider-electric.com/files?p_Doc_Ref=SPD_VAVR-5WKLPK_LS.
- [36] “Pérdida de aislamiento eléctrico.” <https://www.mipodo.com/blog/informacion/perdida-aislamiento-como-identificarla/> (accessed Jul. 06, 2021).
- [37] “Centrosur – PORTAL DE LA CENTROSUR.” <https://www.centrosur.gob.ec/?q=node%2F11> (accessed Jun. 22, 2021).
- [38] CONELEC, “Calidad de servicio eléctrico de distribución,” Accessed: Jul. 14, 2021. [Online]. Available: <https://www.regulacionelectrica.gob.ec/wp-content/uploads/downloads/2016/02/Regulacion-No.-CONELEC-004-01.pdf>.
- [39] J. Manuel Gutiérrez, “Introducción a las Redes Neuronales,” Accessed: Sep. 22, 2021. [Online]. Available: <http://ccaix3.unican.es/~gutierjmhttp://ccaix3.unican.es/~AIGroupJAVA:http://www.aist.go.jp/NIBH/~b0616/Lab/Links.html>.
- [40] M. En *et al.*, ““PREDICCIÓN DE LA DEMANDA ELÉCTRICA: COMPARATIVA ARIMA-REDES NEURONALES MEDIANTE SOFTWARE SPSS’ AUTOR.”
- [41] “Training Neural Network using PyTorch | by Tasnuva Zaman | Towards Data Science.” <https://towardsdatascience.com/training-a-neural-network-using-pytorch-72ab708da210> (accessed Jun. 21, 2021).
- [42] “CONCEPTOS BÁSICOS SOBRE REDES NEURONALES.” Accessed: Dec. 27, 2021. [Online]. Available: <http://grupo.us.es/gtocoma/pid/pid10/RedesNeuronales.htm>.
- [43] P. Larranaga, “Tema 8. Redes Neuronales The correlation between the brain activity and energy metabolism View project Thermal Camera View project,” Accessed: Sep. 29, 2021. [Online]. Available: <https://www.researchgate.net/publication/268291232>.
- [44] “Introduction Perceptrons Training with Delta rule.”
- [45] “Tutorial de redes neuronales recurrentes, parte 1 - Introducción a las RNN - WildML.” <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/> (accessed Sep. 29, 2021).
- [46] “Entendiendo LSTM Networks - blog de colah.” <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> (accessed Sep. 30, 2021).
- [47] “Qué es el aprendizaje automático y hacia dónde se dirige.” <http://www.cs.cmu.edu/~tom/>

(accessed Sep. 30, 2021).

- [48] “El ‘deep learning’ está de moda.” https://www.eldiario.es/hojaderouter/tecnologia/software/moda-deep-learning-algoritmo-inteligencia-artificial_1_4792941.html (accessed Sep. 30, 2021).
- [49] ArcMap, “¿Qué es el geoprocesamiento?” <https://desktop.arcgis.com/es/arcmap/10.3/main/analyze/what-is-geoprocessing.htm>.
- [50] “PyTorch.” <https://pytorch.org/>.
- [51] “Error Porcentual Absoluto Medio (MAPE) en un Pronóstico de Demanda.” <https://www.gestiondeoperaciones.net/proyeccion-de-demanda/error-porcentual-absoluto-medio-mape-en-un-pronostico-de-demanda/> (accessed Sep. 21, 2021).
- [52] S. Maithani, “A neural network based urban growth model of an Indian city,” *J. Indian Soc. Remote Sens.* 2009 373, vol. 37, no. 3, pp. 363–376, Jan. 2010, doi: 10.1007/S12524-009-0041-7.
- [53] “Investigación: Medidas de concordancia: el índice de Kappa Medidas de concordancia: el índice de Kappa,” Accessed: Sep. 21, 2021. [Online]. Available: www.fisterra.com.
- [54] “¿Qué es el coeficiente de correlación de Pearson?” <https://www.questionpro.com/blog/es/coeficiente-de-correlacion-de-pearson/> (accessed Sep. 21, 2021).
- [55] “Regresión lineal en Python.” <https://machinelearningparatodos.com/regresion-lineal-en-python/> (accessed Sep. 13, 2021).
- [56] “sklearn.metrics.recall_score - documentación de scikit-learn 0.24.2.” https://scikit-learn.org/stable/modules/generated/sklearn.metrics.recall_score.html (accessed Sep. 13, 2021).
- [57] GISGeography, “Interpolacion de ponderacion de distancia inversa (IDW),” 2021. <https://gisgeography.com/inverse-distance-weighting-idw-interpolation/>.
- [58] J. Camejo, H. Gonzalez, and C. Morell, “An empirical evaluation of the MLP neural network model for multitarget regression,” *Ser. Científica la Univ. las Ciencias Informáticas*.
- [59] “sklearn.neural_network.MLPRegressor — scikit-learn 0.24.2 documentation.” https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html (accessed Sep. 13, 2021).
- [60] “sklearn.neural_network.MLPClassifier - documentación de scikit-learn 0.24.2.” https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html (accessed Sep. 13, 2021).
- [61] “K-Means con Python paso a paso | Aprende Machine Learning.” <https://www.aprendemachinelearning.com/k-means-en-python-paso-a-paso/> (accessed Sep. 13, 2021).
- [62] M. Peña, “Universidad nacional de piura escuela de posgrado programa de maestría en ingeniería informática,” *Sist. Inf. Mini Estac. Metereológica Para La Captura Datos Temp. Humed. Y Presión Atmosférica - Cent. Operaciones Emerg. Reg. - Piura*, pp. 30–31, 2019.
- [63] “Precipitación - Netatmo.” <https://www.netatmo.com/es-es/glosario/precipitacion>

(accessed Sep. 03, 2021).

- [64] “ EL VIENTO .” <https://www.canaltiempo21.com/capitulo-6-el-viento/#tres> (accessed Sep. 03, 2021).

GLOSARIO

AAM Algoritmos de Aprendizaje de Máquina

ANN Red Neuronal Artificial

APM Aprendizaje Profundo de Máquina

DNN Red Neuronal Profunda

EERCS Empresa Eléctrica Regional CENTRO SUR

GIS Sistema de Información Geográfica

GPU Unidad de Procesador Gráfico

IA Inteligencia Artificial

IDW Ponderación de Distancia Inversa

LBFGS Algoritmo de Memoria Limitada de Broyden-Fletcher-Goldfarb-Shanno

LSM Razón Cuadrada Mínima

MAPE Error Medio Absoluto

MLP Perceptron Multicapa

PMC Perceptron Multicapa

RELU Función Unidad Lineal Rectificada

RMLP Redes de Memoria a Largo Plazo

RMSE Error Cuadrático Medio

RN Red Neuronal

RNR Red Neuronal Recurrente

R^2 Regresión Lineal

SD Sistema de Distribución

SEP Sistema Eléctrico de Potencia

SGD Descenso de Gradiente Estocástico

SIGADE Departamento de Análisis y Sistemas de Distribución

ANEXOS

ANEXO A

Tabla 1 de registro de las interrupciones no programadas generadas durante los años 2015 al 2019 con un registro de información.

- **AÑO 2015**

Tabla 1. Total, de interrupciones no programadas registradas en el año 2015 (Subcausas).

CSBCAUSA	Total, de Interrupciones
ACTUA RELE DE SOBRECORRIENTE	71
ACTUA SISTEMA DE PROTECCIONES	52
AISLADOR ROTO	21
ANIMALES EN RED	103
AREGLO DE PUENTES	7
CAMBIAR BASES DE PORTAFUSIBLES BT	310
CAMBIAR POSTE	80
CAMBIAR SECCIONADOR	31
CAMBIAR TIRAFUSIBLE	93
CAMBIO DE CONECTOR	112
CAMBIO DE FUSIBLE NH	126
CAMBIO LUMINARIA	2
CERRAR SECCIONADOR	6
CHOQUE	164
CONDUCTOR RECOCIDO	45
CONEXION DOLOSA	11
CORTOCIRCUITO	1551
CRUCETA AVERIADA	4
DAÑO INTERNO	3
DAÑO OCASIONADO POR OTRAS EMPR.SERV.PUBL	7
DESCARGAS ATMOSFERICAS	368
DESLAVE DERRUMBE	1
DESLIZAMIENTO DE TIERRA	13

DISYUNTOR DEFECTUOSO	2
FALLA EQUIPO MEDICION (TC TP)	4
FALLA TRANSITORIA	228
FALLAS POR MANIOBRAS	18
INCENDIO	4
INUNDACIÓN	1
INUNDACIONES	1
LÍNEA ARRANCADA	239
MATERIALES LLEVADOS POR EL VIENTO	10
NO DAN RAZON	3
OTRAS	180
PARARRAYO DEFECTUOSO	18
PERSONAL NO AUTORIZADO MANIPULA RED	39
PUENTE EN MT ARRANCADO	75
RAMAS SOBRE LA RED	756
RED FLOJA	166
SECCIONADOR AVERIADO	55
TEMPESTAD O TORMENTA	33
TENSOR SUELTO	7
TRAFO O RED SOBRECARGADA	87
TRANSFORMADOR QUEMADO	69
VIENTO	192
Total, general	5368

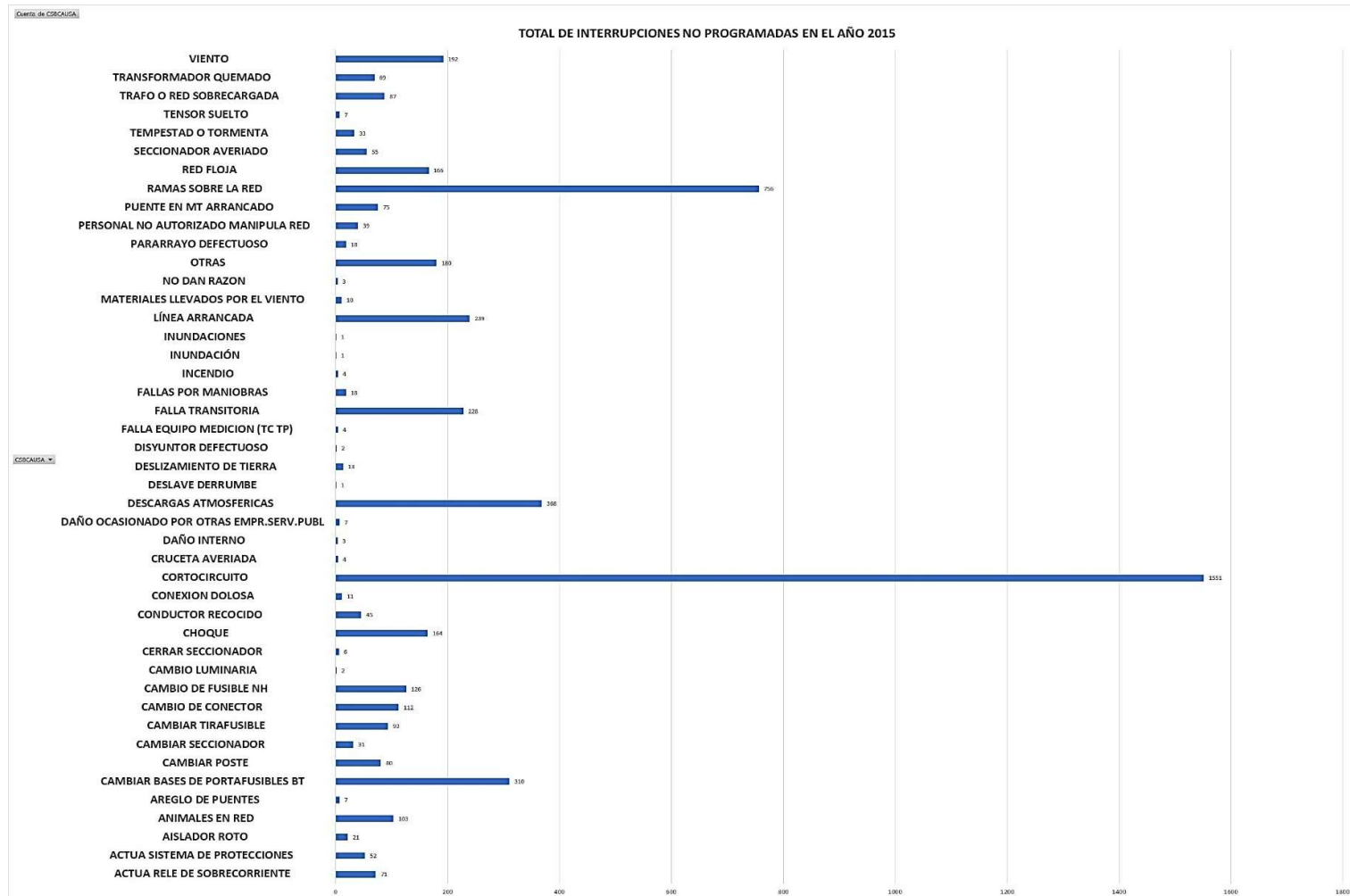


Figura 30. Representación de las interrupciones en el año 2015.

- **AÑO 2016**

Tabla 2. Total, de interrupciones no programadas registradas en el año 2016 (Subcausas).

CSBCAUSA	Total, de Interrupciones
ACTUA RELE DE SOBRECORRIENTE	3
ACTUA SISTEMA DE PROTECCIONES	68
AISLADOR ROTO	22
ANIMALES EN RED	149
AREGLO DE PUENTES	13
CAMBIAR BASES DE PORTAFUSIBLES BT	244
CAMBIAR POSTE	77
CAMBIAR SECCIONADOR	38
CAMBIAR TIRAFUSIBLE	219
CAMBIO DE CONECTOR	169
CAMBIO DE FUSIBLE NH	391
CERRAR SECCIONADOR	2
CHOQUE	181
CONDUCTOR RECOCIDO	38
CONEXION DOLOSA	4
CORTOCIRCUITO	1463
CRUCETA AVERIADA	2
DAÑO INTERNO	1
DAÑO OCASIONADO POR OTRAS EMPR.SERV.PUBL	3
DESCARGAS ATMOSFERICAS	574
DESLAVE DERRUMBE	2
DESLIZAMIENTO DE TIERRA	9
FALLA EQUIPO MEDICION (TC TP)	5
FALLA TRANSITORIA	413
FALLAS POR MANIOBRAS	21
INCENDIO	3

INUNDACIONES	3
LÍNEA ARRANCADA	400
MATERIALES LLEVADOS POR EL VIENTO	5
OTRAS	68
PARARRAYO DEFECTUOSO	9
PERSONAL NO AUTORIZADO MANIPULA RED	17
PUENTE EN MT ARRANCADO	35
RAMAS SOBRE LA RED	999
RED FLOJA	204
SECCIONADOR AVERIADO	47
TEMPESTAD O TORMENTA	36
TENSOR SUELTO	73
TERREMOTO	6
TRAFO O RED SOBRECARGADA	75
TRANSFORMADOR QUEMADO	66
VIENTO	197
Total general	6354

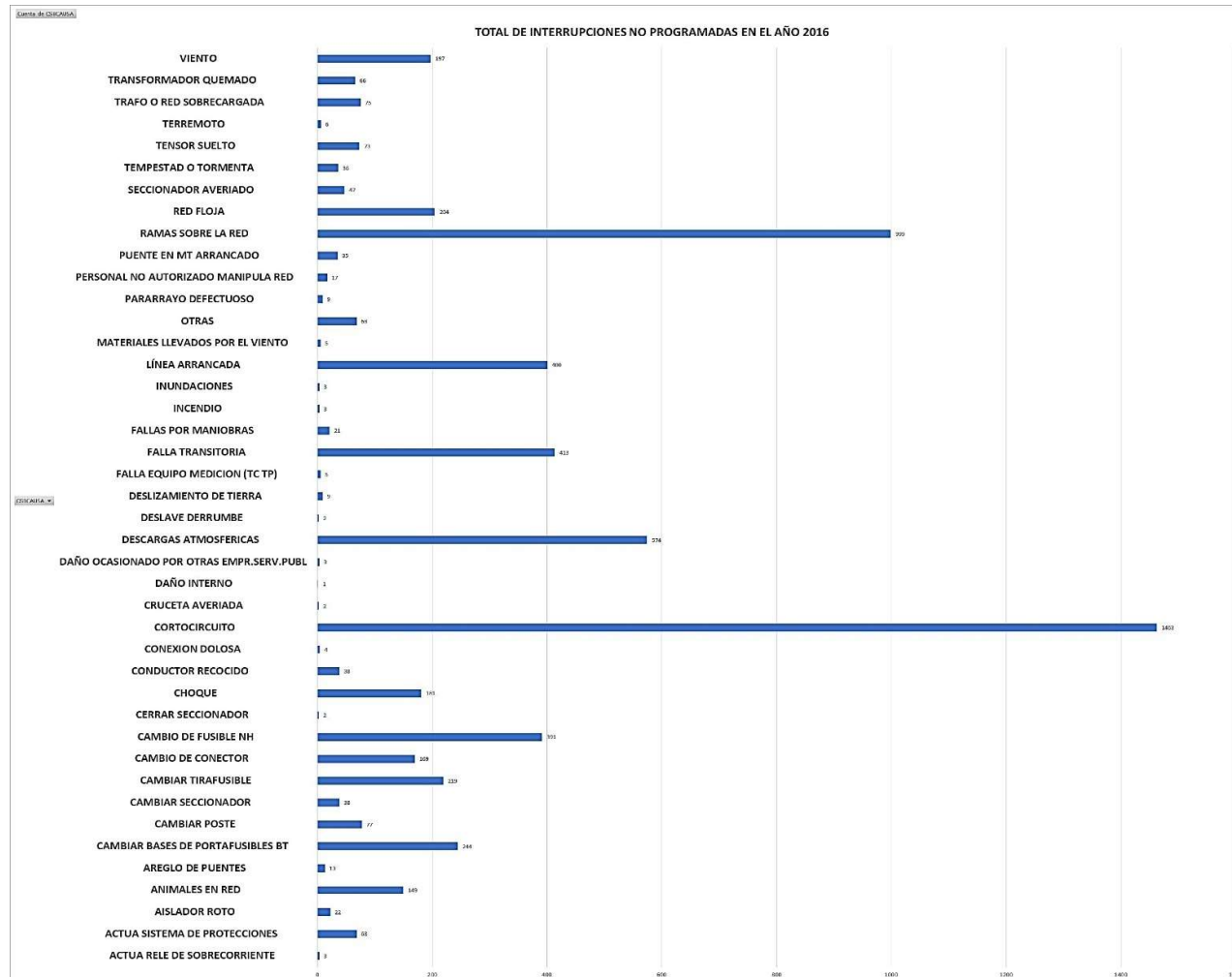


Figura 31. Representación de las interrupciones en el año 2016.

- **AÑO 2017**

Tabla 3. Total, de interrupciones no programadas registradas en el año 2017 (Subcausas).

CSBCAUSA	Total, de Interrupciones
ACTUA RELE DE SOBRECORRIENTE	4
ACTUA SISTEMA DE PROTECCIONES	52
AISLADOR ROTO	27
ANIMALES EN RED	170
AREGLO DE PUENTES	10
CAMBIAR BASES DE PORTAFUSIBLES BT	290
CAMBIAR POSTE	103
CAMBIAR SECCIONADOR	56
CAMBIAR TIRAFUSIBLE	198
CAMBIO DE CONECTOR	149
CAMBIO DE FUSIBLE NH	301
CAMBIO LUMINARIA	2
CERRAR SECCIONADOR	23
CHOQUE	167
CONDUCTOR RECOCIDO	33
CONEXION DOLOSA	4
CONTAMINACION	1
CORTOCIRCUITO	1429
CRUCETA AVERIADA	5
DAÑO INTERNO	2
DAÑO OCASIONADO POR OTRAS EMPR.SERV.PUBL	5
DESCARGAS ATMOSFERICAS	693
DESLAVE DERRUMBE	11
DESLIZAMIENTO DE TIERRA	66
FALLA TRANSITORIA	350
FALLAS POR MANIOBRAS	5
INCENDIO	5
INUNDACIÓN	1

INUNDACIONES	8
LÍNEA ARRANCADA	381
MATERIALES LLEVADOS POR EL VIENTO	4
OTRAS	40
PARARRAYO DEFECTUOSO	10
PERSONAL NO AUTORIZADO MANIPULA RED	26
PUENTE EN MT ARRANCADO	64
RAMAS SOBRE LA RED	1077
RED FLOJA	275
SECCIONADOR AVERIADO	64
TEMPESTAD O TORMENTA	46
TENSOR SUELTO	65
TRAFO O RED SOBRECARGADA	66
TRANSFORMADOR QUEMADO	88
VIENTO	196
(en blanco)	
Total general	6572

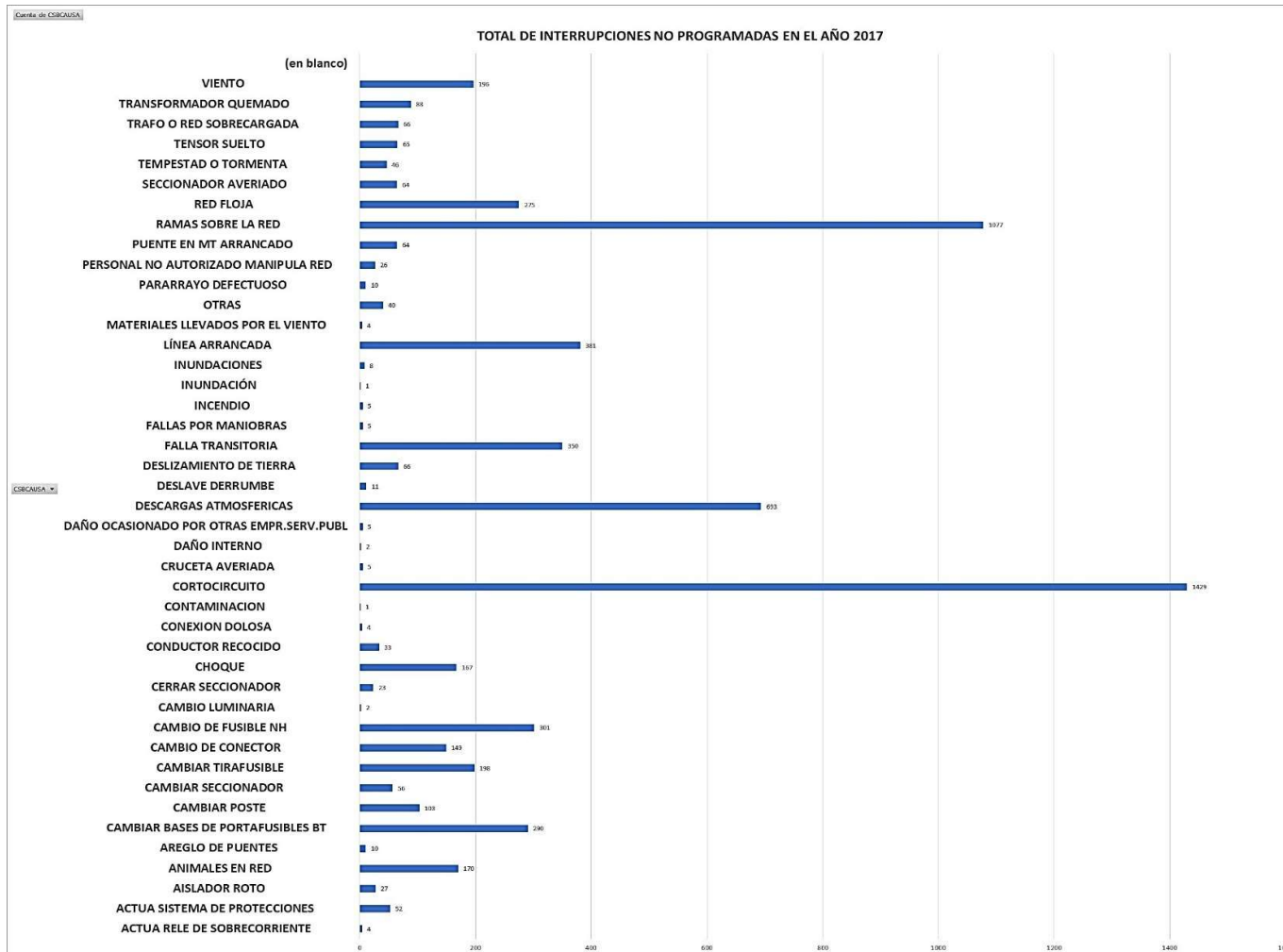


Figura 32 Representación de las interrupciones en el año 2017.

- **AÑO 2018**

Tabla 4. Total, de interrupciones no programadas registradas en el año 2018 (Subcausas).

CSBCAUSA	Total, de Interrupciones
ACTUA RELE DE SOBRECORRIENTE	2
ACTUA SISTEMA DE PROTECCIONES	60
AISLADOR ROTO	11
ANIMALES EN RED	113
AREGLO DE PUENTES	14
CAMBIAR BASES DE PORTAFUSIBLES BT	116
CAMBIAR POSTE	34
CAMBIAR SECCIONADOR	33
CAMBIAR TIRAFUSIBLE	249
CAMBIO DE CONECTOR	110
CAMBIO DE FUSIBLE NH	311
CERRAR SECCIONADOR	8
CHOQUE	177
CONDUCTOR RECOCIDO	18
CONEXION DOLOSA	4
CONTAMINACION	29
CORTOCIRCUITO	1205
CRUCETA AVERIADA	1
DAÑO INTERNO	11
DAÑO OCASIONADO POR OTRAS EMPR.SERV.PUBL	88
DESCARGAS ATMOSFERICAS	643
DESLAVE DERRUMBE	3
DESLIZAMIENTO DE TIERRA	31
FALLA EQUIPO MEDICION (TC TP)	1
FALLA TRANSITORIA	2458
FALLAS POR MANIOBRAS	9

INCENDIO	7
INUNDACIÓN	3
INUNDACIONES	1
LÍNEA ARRANCADA	189
MATERIALES LLEVADOS POR EL VIENTO	94
OTRAS	378
PARARRAYO DEFECTUOSO	5
PERSONAL NO AUTORIZADO MANIPULA RED	156
PUENTE EN MT ARRANCADO	32
RAMAS SOBRE LA RED	965
RED FLOJA	233
SECCIONADOR AVERIADO	32
TEMPESTAD O TORMENTA	49
TENSOR SUELTO	65
TERREMOTO	4
TRAFO O RED SOBRECARGADA	37
TRANSFORMADOR QUEMADO	40
VIENTO	174
(en blanco)	
Total general	8203

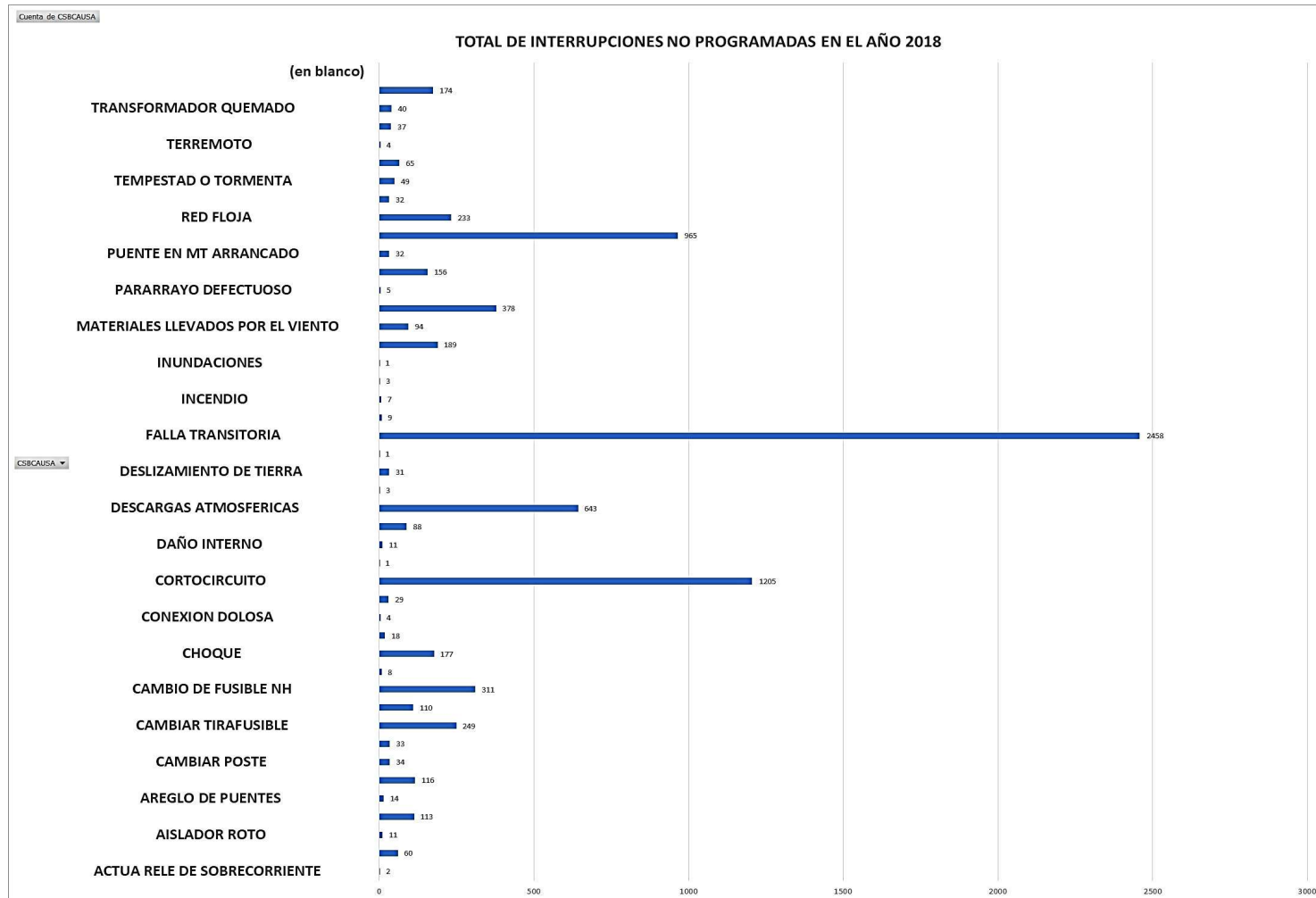


Figura 33. Representación de las interrupciones en el año 2018.

- **AÑO 2019**

Tabla 5. Total, de interrupciones no programadas registradas en el año 2019 (Subcausas).

CSBCAUSA	Total, de Interrupciones
ACTUA SISTEMA DE PROTECCIONES	76
AISLADOR ROTO	5
ANIMALES EN RED	108
AREGLO DE PUENTES	5
CAMBIAR BASES DE PORTAFUSIBLES BT	77
CAMBIAR POSTE	32
CAMBIAR SECCIONADOR	13
CAMBIAR TIRAFUSIBLE	358
CAMBIO DE CONECTOR	100
CAMBIO DE FUSIBLE NH	372
CERRAR SECCIONADOR	4
CHOQUE	131
CONDUCTOR RECOCIDO	3
CONTAMINACION	9
CORTOCIRCUITO	980
CRUCETA AVERIADA	1
DAÑO INTERNO	15
DAÑO OCASIONADO POR OTRAS EMPR.SERV.PUBL	120
DESCARGAS ATMOSFERICAS	450
DESLAVE DERRUMBE	2
DESLIZAMIENTO DE TIERRA	23
FALLA EQUIPO MEDICION (TC TP)	4
FALLA TRANSITORIA	2390
FALLAS POR MANIOBRAS	9
INCENDIO	1
INUNDACIÓN	1
INUNDACIONES	4

LÍNEA ARRANCADA	158
MATERIALES LLEVADOS POR EL VIENTO	69
NO HAY ACCESO	2
OTRAS	267
PARARRAYO DEFECTUOSO	2
PERSONAL NO AUTORIZADO MANIPULA RED	34
PUENTE EN MT ARRANCADO	13
RAMAS SOBRE LA RED	868
RECONECTADOR DEFECTUOSO	1
RED FLOJA	202
SECCIONADOR AVERIADO	29
TEMPESTAD O TORMENTA	24
TENSOR SUELTO	86
TERREMOTO	16
TRAFO O RED SOBRECARGADA	14
TRANSFORMADOR QUEMADO	19
VIENTO	137
(en blanco)	
Total, general	7234

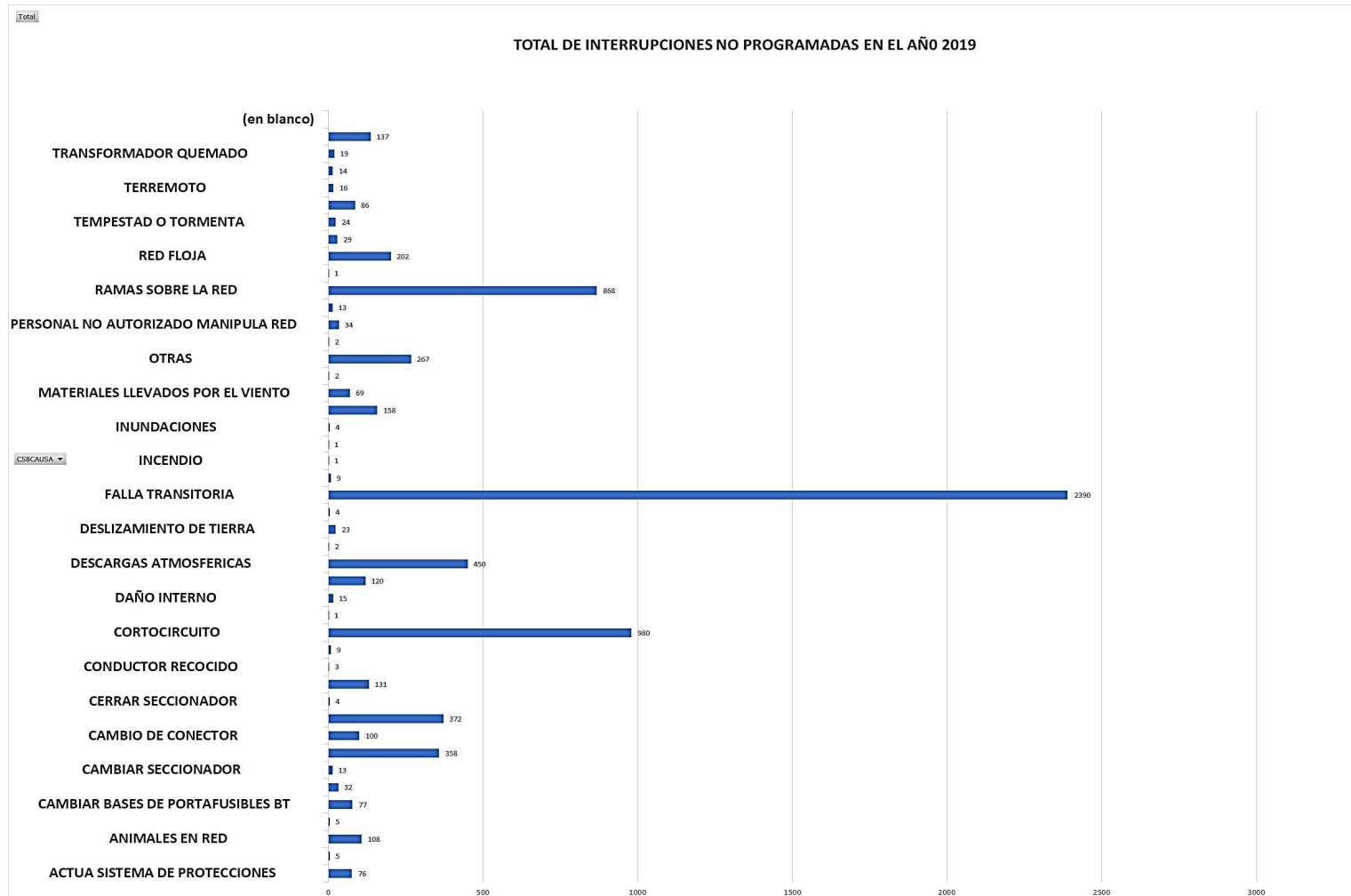


Figura 34 Representación de las interrupciones en el año 2019.

ANEXO B

Tabla 6 de registro de las interrupciones de la CENTROSUR con las entradas que se trabajara para la red neuronal (se presenta un registro de 33732 interrupciones no programadas).

Tabla 6. Registro de interrupciones de la Empresa Eléctrica.

INTALI	CSPCAUSA	CSBCAUSA	I3ECAUSA	I3EBCAUSA	INTCORXALI	INTCORYALI	INTFEI	INTFEX	AÑO	MESES	DIAS	DURACION
1221	Internas No Programadas	CAMBIAR BASES DE PORTAFUSIBLES BT	Equipos	Estructura/Apoyo	733000,831	9684465,857	1/1/2015 14:42	1/1/2015 14:53	2015	1	1	0,18333333
526	Internas No Programadas	CHOQUE	Público	Choque de vehiculo	717972,604	9678170,132	2/1/2015 17:57	2/1/2015 19:42	2015	1	2	1,75
527	Internas No Programadas	CHOQUE	Público	Choque de vehiculo	717974,707	9678169,408	2/1/2015 17:57	2/1/2015 20:55	2015	1	2	2,96666667
526	Internas No Programadas	CHOQUE	Público	Choque de vehiculo	718331,12	9678270,284	2/1/2015 21:21	3/1/2015 18:00	2015	1	2	20,65
523	Internas No Programadas	CAMBIAR BASES DE PORTAFUSIBLES BT	Equipos	Estructura/Apoyo	719404,687	9680457,96	3/1/2015 7:53	3/1/2015 8:03	2015	1	3	0,16666667
523	Internas No Programadas	LÍNEA ARRANCADA	Equipos	Conductor/Cable	718395,3	9682728,71	3/1/2015 10:52	3/1/2015 11:15	2015	1	3	0,38333333
527	Internas No Programadas	LÍNEA ARRANCADA	Equipos	Conductor/Cable	713768,071	9682391,824	2/1/2015 16:35	3/1/2015 11:06	2015	1	2	18,51666667
325	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	724116,998	9680248,256	3/1/2015 13:18	3/1/2015 14:36	2015	1	3	1,3
521	Internas No Programadas	CHOQUE	Público	Choque de vehiculo	719741,737	9673030,365	4/1/2015 8:21	4/1/2015 13:23	2015	1	4	5,03333333
325	Internas No Programadas	ANIMALES EN RED	Fauna silvestre	Otro mamifero	724116,998	9680248,256	4/1/2015 14:30	4/1/2015 15:07	2015	1	4	0,61666667
721	Internas No Programadas	TRANSFORMADOR QUEMADO	Desconocidas	Causa no especificada	726667,897	9681526,605	4/1/2015 22:00	5/1/2015 13:18	2015	1	4	15,3
521	Internas No Programadas	PUENTE EN MT ARRANCADO	Equipos	Conector/Empalme	730941,984	9647892,93	5/1/2015 11:47	5/1/2015 12:48	2015	1	5	1,01666667
522	Internas No Programadas	CHOQUE	Público	Choque de vehiculo	719672,239	9679160,579	5/1/2015 12:08	5/1/2015 13:17	2015	1	5	1,15
721	Internas No Programadas	RAMAS SOBRE LA RED	Vegetación	Arbol/Rama	725482,807	9683240,508	5/1/2015 11:45	5/1/2015 12:23	2015	1	5	0,63333333
721	Internas No Programadas	RAMAS SOBRE LA RED	Vegetación	Arbol/Rama	724493,249	9684447,32	5/1/2015 12:23	5/1/2015 13:36	2015	1	5	1,21666667
721	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	724493,249	9684447,32	5/1/2015 15:55	5/1/2015 16:45	2015	1	5	0,83333333

1424	Internas No Programadas	RED FLOJA	Equipos	Conductor/Cable	691550,19	9640955,348	5/1/2015 15:30	5/1/2015 17:00	201 5	1	5	1,5
723	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	726671,019	9692618,112	5/1/2015 19:33	5/1/2015 21:05	201 5	1	5	1,5333333 33
325	Internas No Programadas	CAMBIAR BASES DE PORTAFUSIBLES BT	Equipos	Estructura/Apoyo	725086,721	9680213,36	5/1/2015 21:44	5/1/2015 22:01	201 5	1	5	0,2833333 33
722	Internas No Programadas	LÍNEA ARRANCADA	Equipos	Conductor/Cable	727120,918	9680754,145	5/1/2015 22:15	5/1/2015 22:54	201 5	1	5	0,65
1521	Internas No Programadas	RAMAS SOBRE LA RED	Vegetación	Arbol/Rama	750096,233	9677502,743	6/1/2015 11:38	6/1/2015 13:20	201 5	1	6	1,7
1523	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	748621,028	9680371,219	1/1/2015 22:29	2/1/2015 8:15	201 5	1	1	9,7666666 67
1523	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	747310,918	9680685,447	2/1/2015 13:36	2/1/2015 16:20	201 5	1	2	2,7333333 33
1522	Internas No Programadas	CHOQUE	Público	Choque de vehiculo	745907,288	9680424,204	3/1/2015 13:18	3/1/2015 14:20	201 5	1	3	1,0333333 33
1522	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	745196,478	9682920,326	3/1/2015 14:31	3/1/2015 17:00	201 5	1	3	2,4833333 33
1521	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	748983,176	9669798,858	4/1/2015 11:29	4/1/2015 12:10	201 5	1	4	0,6833333 33
1521	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	742800,571	9674017,181	4/1/2015 8:18	4/1/2015 9:50	201 5	1	4	1,5333333 33
1521	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	745448,965	9674926,818	5/1/2015 12:30	5/1/2015 17:00	201 5	1	5	4,5
523	Internas No Programadas	LÍNEA ARRANCADA	Equipos	Conductor/Cable	720952,037	9685869,535	6/1/2015 17:00	6/1/2015 17:20	201 5	1	6	0,3333333 33
1424	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	669148,29	9634621,296	6/1/2015 22:25	7/1/2015 17:20	201 5	1	6	#iVALOR !
1421	Internas No Programadas	AISLADOR ROTO	Equipos	Aislador	673532,872	9627291,803	6/1/2015 22:27	7/1/2015 13:00	201 5	1	6	14,55
1424	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	669148,29	9634621,296	6/1/2015 22:30	8/1/2015 13:00	201 5	1	6	#iVALOR !
723	Internas No Programadas	RAMAS SOBRE LA RED	Vegetación	Arbol/Rama	723809,006	9685167,302	7/1/2015 9:30	7/1/2015 18:57	201 5	1	7	9,45
1521	Internas No Programadas	CORTOCIRCUITO	Desconocidas	Causa no especificada	746876,689	9675591,438	6/1/2015 17:12	6/1/2015 18:10	201 5	1	6	0,9666666 67
821	Internas No Programadas	RED FLOJA	Equipos	Conductor/Cable	722823,943	9677256,374	7/1/2015 20:44	7/1/2015 21:19	201 5	1	7	0,5833333 33
721	Internas No Programadas	CAMBIO DE CONECTOR	Equipos	Conector/Empalme	724931,011	9682691,485	8/1/2015 9:12	8/1/2015 9:41	201 5	1	8	0,4833333 33

ANEXO C

Mapas de calor de las variables externas interpoladas.

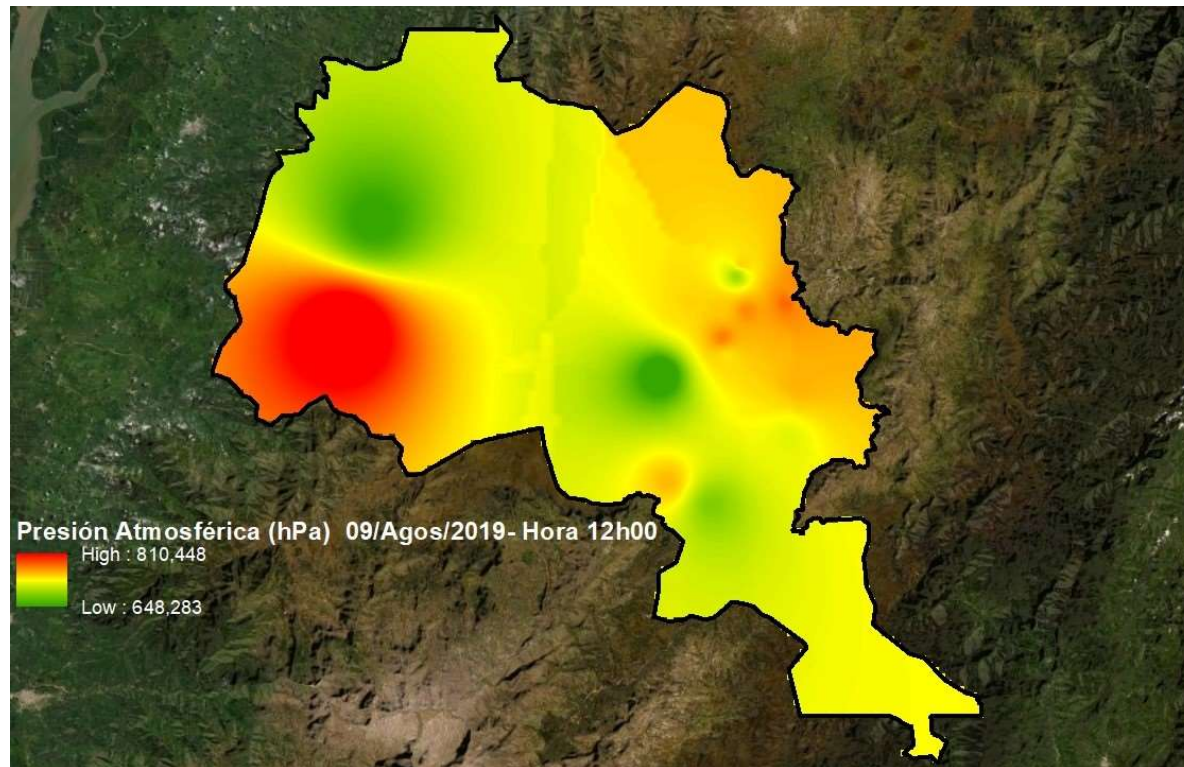


Figura 35. Mapa de calor de la Presión Atmosférica.

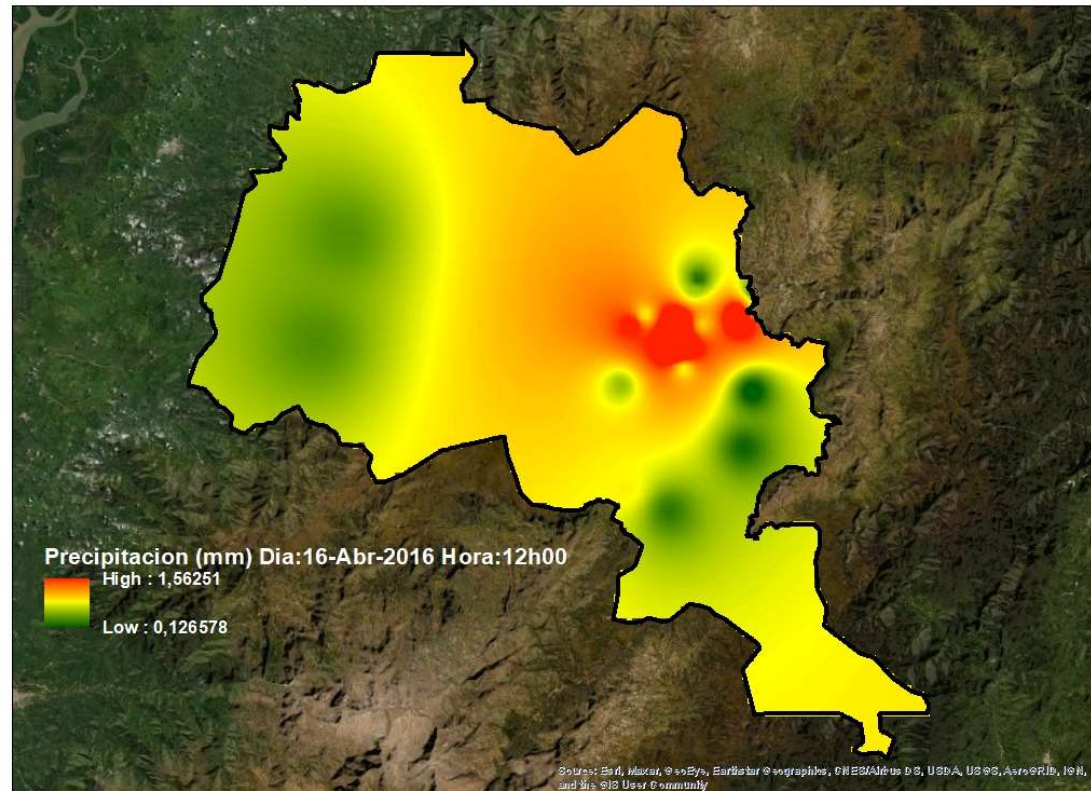


Figura 36. Mapa de Calor de la Precipitación.

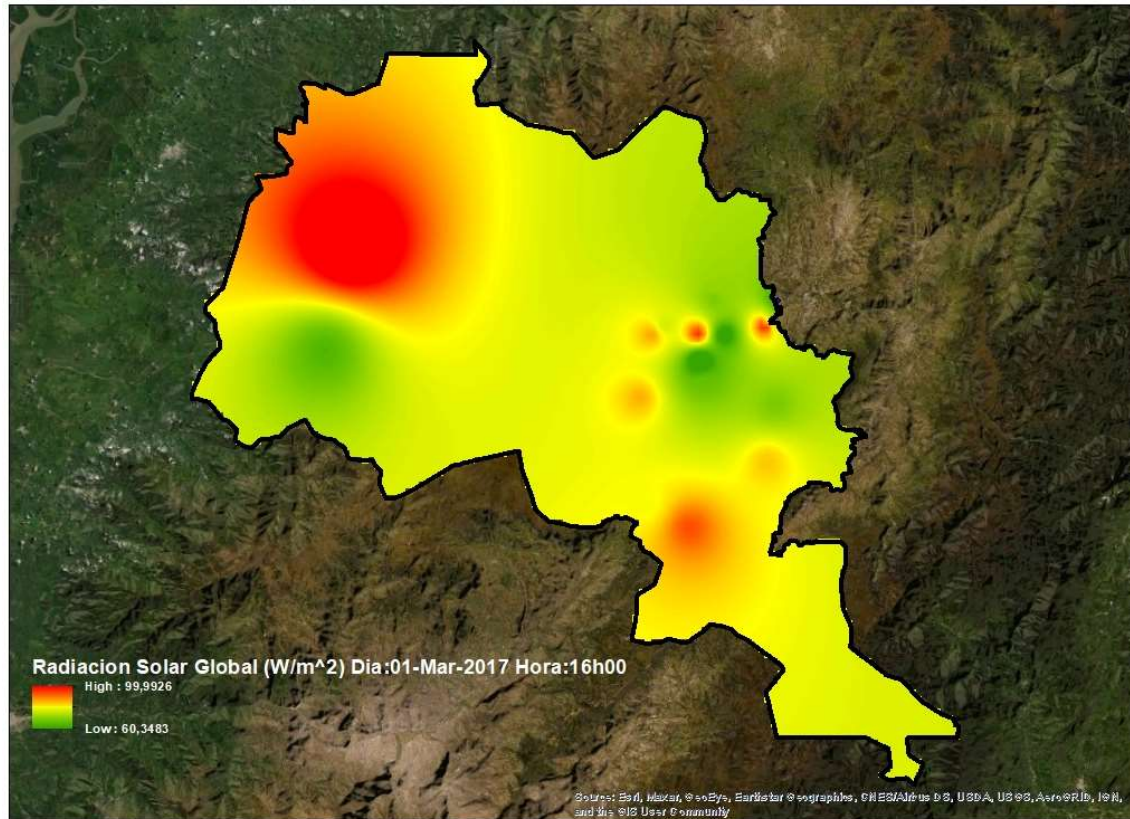


Figura 37. Mapa de Calor de la Radiación Solar.

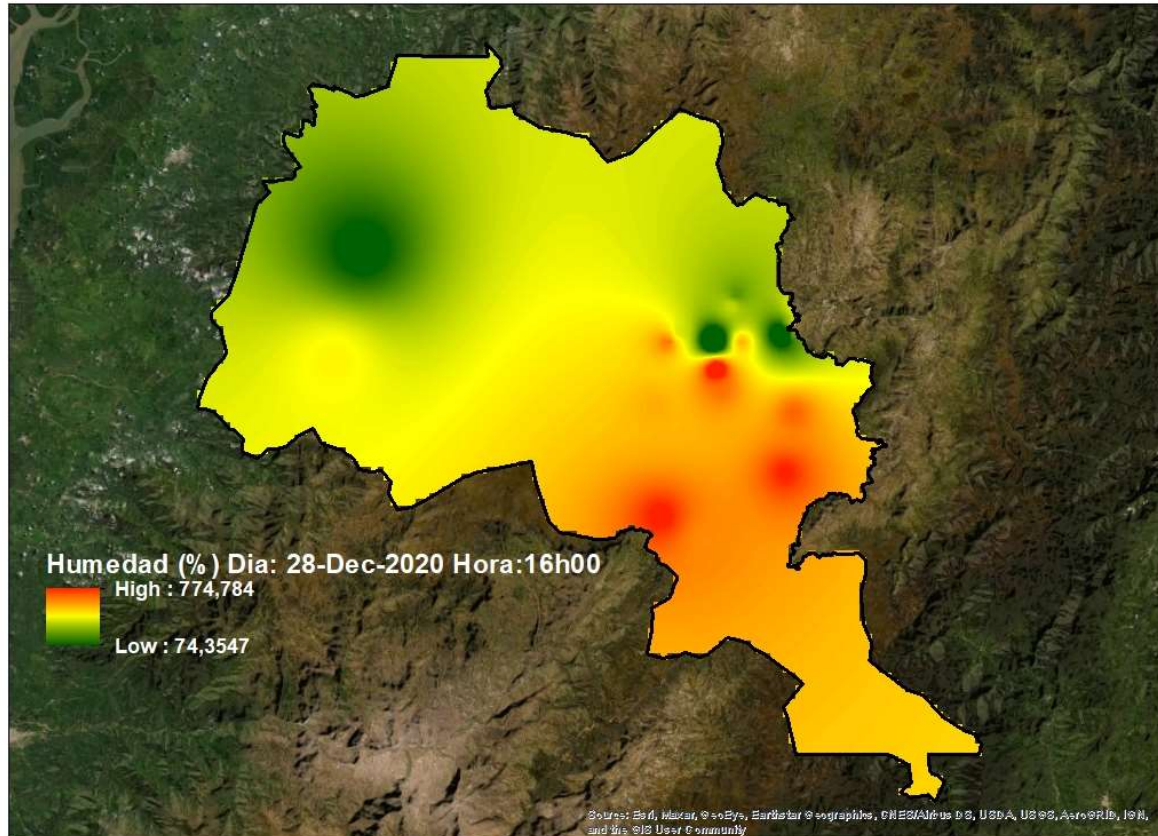


Figura 38. Mapa de calor de la Humedad.

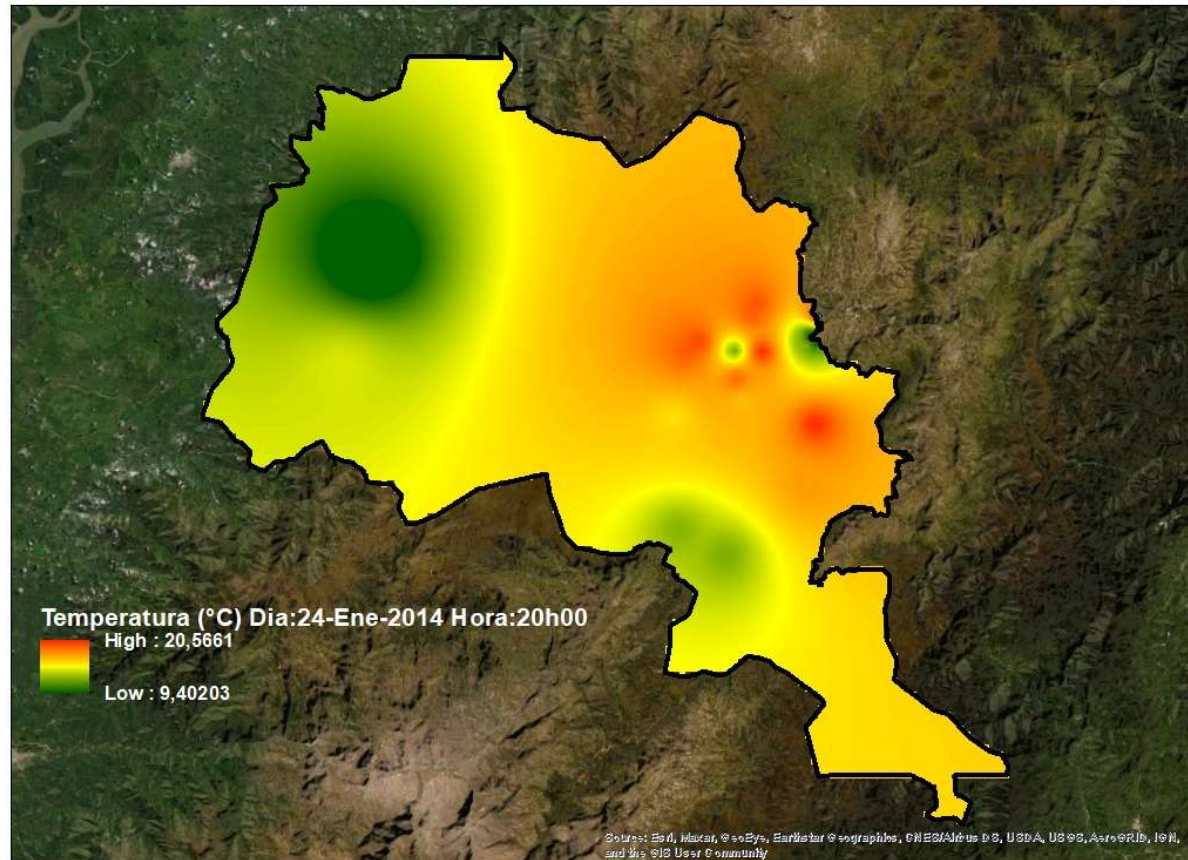


Figura 39. Mapa de calor de la Temperatura.

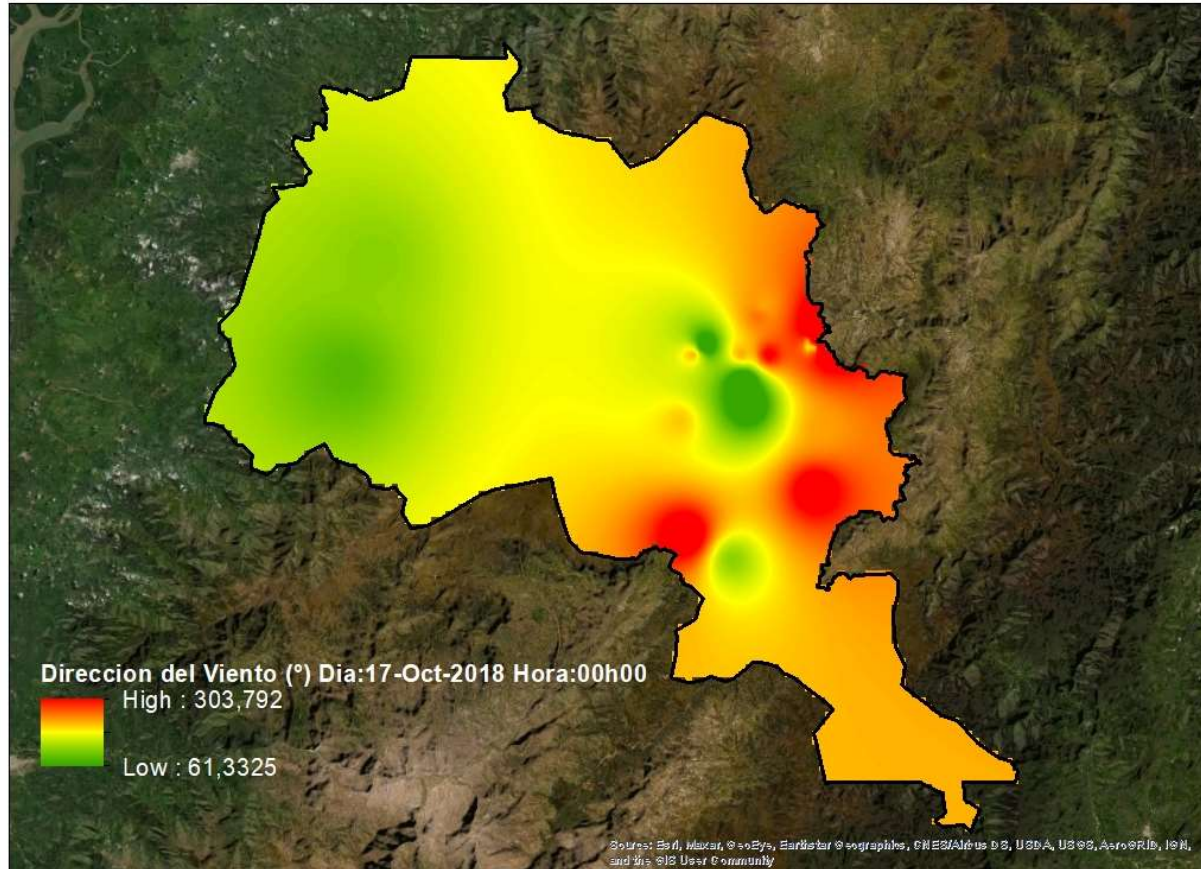


Figura 40. Mapa de calor de la Dirección del Viento.

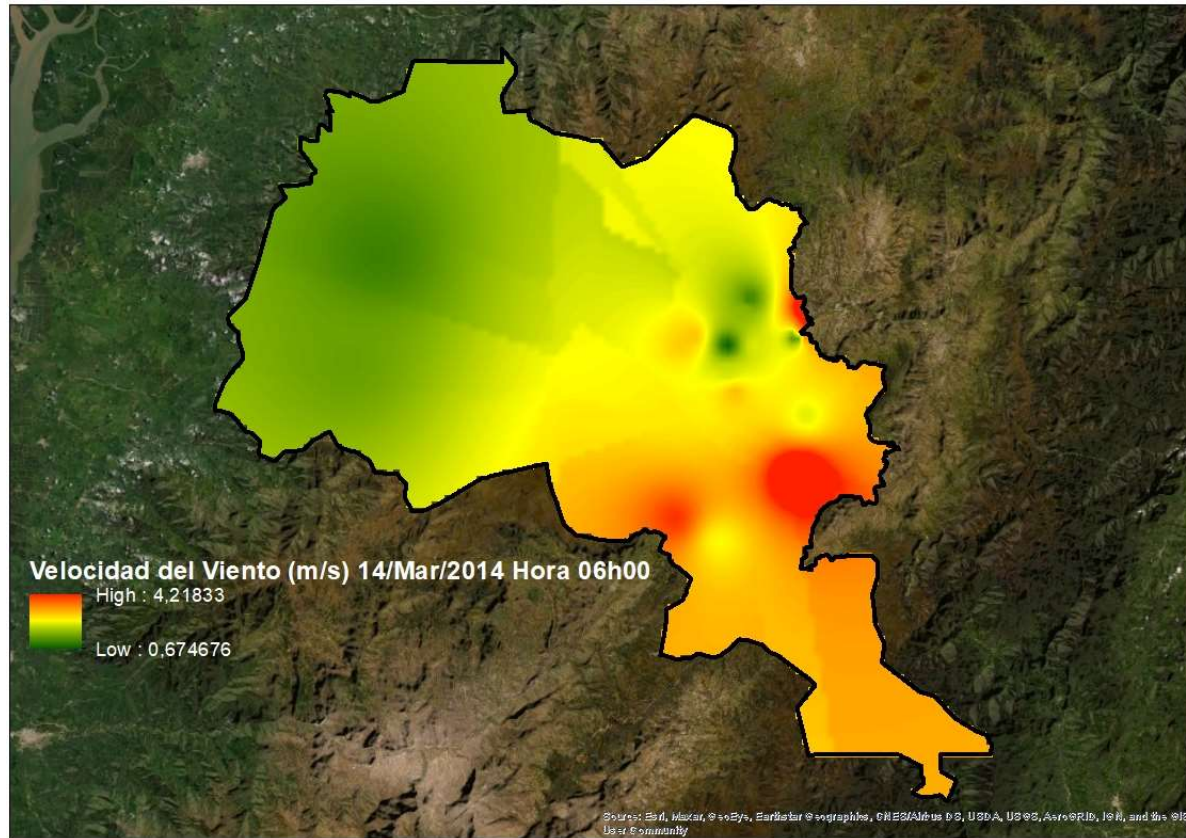


Figura 41. Mapa de calor de la Velocidad del Viento.