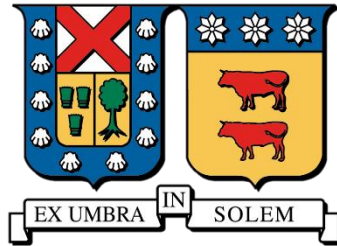


Universidad técnica Federico Santa María

Departamento de ingeniería eléctrica

Santiago, Chile



EVALUACIÓN DE LA CAPACIDAD DE PREDICCIÓN DE LA EVOLUCIÓN DE ÁRBOLES ELÉCTRICOS EN RESINA EPÓXICA MEDIANTE EL USO DE TÉCNICAS DE INTELIGENCIA ARTIFICIAL

Estudiante

Franco Salas Henríquez

Profesor guía:

Ricardo Álvarez

Correferente:

Roger Schurch

Santiago, 8/9/2023

Agradecimientos.

Agradezco principalmente a mi familia Ivonne, Débora y Carlos, que siempre estuvieron en los momentos difíciles apoyándome incondicionalmente forjándome como persona y permitiéndome llegar a esta instancia en mi desarrollo. A mi amigo Andrés Collado. Y finalmente a mis amigos y compañeros Carlos Martínez, Fabian Martínez, Ismael Lobos, Tomas Manosalva, Cynthia Vargas y Matías Araya por las experiencias compartidas y el apoyo mutuo a lo largo de la carrera.

Abstract

Failures in electrical equipment due to insulation deterioration lead to interruption and restoration times in the operation of the electrical system, resulting in associated economic losses. One of the aging mechanisms in insulation are electrical trees, cavities in the insulation that grow in shapes similar to botanical trees. Monitoring electrical variables associated with electrical trees, such as partial discharges, would enable preventive actions against potential failures in electrical equipment. In this study, the predictive capability of the evolution of electrical trees was evaluated using artificial intelligence techniques. Data from partial discharges and lengths of electrical trees were collected through tests at 15 kV AC, 15 kV AC + 15 kV DC, and 15 kV AC-15 kV DC in epoxy resin samples. Prediction was performed using random forest and neural network models. The data were labeled with the deterioration state of each electrical tree based on its advancement, determined by inspecting its lengths and partial discharge characteristics. The training data were based on parameters from the IEC 60270 standard and phase-resolved partial discharge patterns. Prediction capabilities for the prediction of the states of electrical trees were achieved with approximately 90% accuracy for both random forest and neural network models. Additionally, it was observed that the phase-resolved partial discharge analysis within the feature set is of greater importance than the parameters of IEC 60270.

Resumen

Los fallos en equipamientos eléctricos debido a deterioros de aislantes producen tiempos de corte y reposición en la operación del sistema eléctrico produciendo pérdidas económicas asociadas. Uno de los mecanismos de envejecimiento en aislantes son los árboles eléctricos, cavidades en el aislante que crecen con formas similares a arboles botánicos. Un monitoreo de las variables eléctricas asociadas a los árboles eléctricos, tales como las descargas parciales, permitiría acciones preventivas ante la posibilidad de una falla en el equipamiento eléctrico. En este trabajo de memoria se evaluó la capacidad de predicción de la evolución de árboles eléctricos mediante técnicas de inteligencia artificial. Se utilizaron datos de las descargas parciales y longitudes de árboles eléctricos ensayados a 15 kV AC, 15kV AC + 15 kV DC y 15kV AC-15kV DC, en probetas de resina epóxica. Para realizar la predicción se utilizaron modelos de random forest y redes neuronales. Los datos se etiquetaron asignando un estado de deterioro de cada árbol eléctrico respecto a su avance basado en la inspección de sus longitudes y características de descargas parciales. Los datos de entrenamiento utilizados fueron basados en parámetros de la norma IEC 60270 y mediante el análisis de diagramas resueltos en fase de las descargas parciales. Se obtuvo una capacidad de predecir los estados de árboles eléctricos con una exactitud de alrededor del 90% con los modelos de random forest y redes neuronales. También se observó que dentro del set de características es de mayor importancia la obtención de los análisis resueltos en fase respecto a los parámetros de la IEC 60270.

Índice de contenidos

Agradecimientos.....	1
Abstract	2
Resumen.....	3
1. Introducción	7
1.1. Justificación del tema.....	7
2. Marco teórico.....	10
2.1. Envejecimiento de aislantes.....	10
2.1.1. Descargas Parciales	11
2.1.2. Árboles eléctricos	11
2.1.3. Características relevantes	13
2.1.4. Técnicas de identificación de características relevantes.	14
2.2. Técnicas de inteligencia artificial	15
2.2.1. Técnicas de aprendizaje	15
2.2.1.1 Aprendizaje supervisado	16
2.2.1.2 Aprendizaje no supervisado	16
2.2.1.3 Aprendizaje reforzado	16
2.2.2. Tipos de modelos	17
2.2.2.1 Árboles de decisión	17
2.2.2.2 Random forest	18
2.2.2.3 Redes Neuronales.....	18
2.3. Estado del arte de técnicas de inteligencia artificial aplicada a predicción de envejecimiento de aislantes.....	21
3. Procedimiento experimental	25
3.1. Probetas de ensayo	25
3.2. Ruido de medición.....	26
3.3. Datos disponibles	26
4. Modelos y Resultados de predicción	27
4.1. Modelos propuestos	27
4.1.1. Características relevantes	29
4.1.2. Selección de características.....	30
4.1.3. Etiquetas de estado.....	30
4.2. Resultados random forest.....	32
4.3. Resultados redes neuronales	35

4.3.1.	Predicción	35
4.3.2.	Ajuste de hiperparámetros red neuronal	40
4.4.	Selección de características.....	41
5.	Conclusiones.....	45
5.1.	Trabajo a futuro	46
6.	Referencias.....	47

Índice de Figuras

Figura 2.1:	Descarga interna	11
Figura 2.2:	Descarga superficial.....	11
Figura 2.3:	Descarga corona.....	11
Figura 2.4:	Árbol eléctrico en resina epóxica [9]	12
Figura 2.5:	Grupos inteligencia artificial	16
Figura 2.6:	Esquema de respuesta de un perceptrón ([19]).....	19
Figura 2.7:	Esquema red neuronal densa feed-forward de una capa oculta ([20]).....	20
Figura 2.8:	Resultado predicción de estado de salud de cables XLPE (fuente [26])	22
Figura 2.9:	Modelos de falla para árbol eléctrico utilizado en [28]	23
Figura 3.1:	Esquema muestra (fuente [4]).....	25
Figura 4.1:	Esquema red neuronal modelo 1 generado por TensorFlow	27
Figura 4.2:	Esquema red neuronal modelo 2 generado por TensorFlow	28
Figura 4.3:	Etapas arboles eléctricos usada en [4]	31
Figura 4.4:	Etapas propagación árbol eléctrico usada en [4]	31
Figura 4.5:	Matriz de confusión para random forest para características IEC	33
Figura 4.6:	Matriz de confusión para random forest para características PRPD.....	33
Figura 4.7:	Matriz de confusión para random forest entrenado con características combinadas IEC y PRPD	34
Figura 4.8:	Importancia de características para random forest entrenado con características combinadas PRPD y IEC	35
Figura 4.9:	Matriz de confusión para red neuronal modelo 1 entrenada con todas las características	38
Figura 4.10:	Matriz de confusión para red neuronal modelo 2 entrenada con todas las características	38
Figura 4.11:	Épocas red neuronal modelo 1 para datos de entrenamiento y validación IEC+PRPD	39

Figura 4.12: Épocas red neuronal modelo 2 para datos de entrenamiento y validación IEC+PRPD	39
Figura 4.13: importancia de características red neuronal modelo 2	40
Figura 4.14: eliminación recursiva de características para random forest	43
Figura 4.15: importancia para el caso de 15 características de la eliminación recursiva	43
Figura 4.16: eliminación recursiva de características para red neuronal modelo 2.....	44
Figura 4.17: importancias para el caso de 14 características de eliminación recursiva.	44

Índice de Tablas

Tabla 1: Factores de envejecimiento en aislantes.....	11
Tabla 2: Ejemplo de datos de motores sumergibles utilizados en [27]	22
Tabla 3: Evaluación exactitud modelo random forest para los tipos de características.....	32
Tabla 4: Exactitud para modelos de red neuronal según tipo de características.....	37
Tabla 5: optimización de hiperparámetros red neuronal	41

Capítulo 1

Introducción

1.1. Justificación del tema

En la actualidad, el uso de aislantes poliméricos está ampliamente extendido en los sistemas eléctricos de alta tensión, principalmente en cables subterráneos y submarinos. La confiabilidad de los sistemas de aislación es fundamental para el funcionamiento del sistema eléctrico, ya que una falla en la aislación puede traer importantes consecuencias como, por ejemplo, la interrupción en el suministro eléctrico y/o el daño irreparable en los equipos eléctricos. Esto, a su vez, conlleva altos tiempos de reparación y de reposición de servicio.

La operación de los sistemas eléctricos requiere sistemas de aislación confiables para su correcta operación. Sin embargo, estos sistemas están habitualmente sometidos a deterioro. Los factores de envejecimiento que producen este deterioro son de 4 tipos: térmicos, eléctricos, ambientales y mecánicos. Estos factores se ven afectados por las tensiones y corrientes producto de la operación de los sistemas eléctricos y/o a factores ambientales externos, los cuales activan distintos mecanismos de envejecimiento [1]. Uno de los principales mecanismos de envejecimiento en aislantes poliméricos son los árboles eléctricos. Los árboles eléctricos son cavidades micro tubulares con geometrías complejas, similares a árboles botánicos, que pueden ocurrir en imperfecciones tales como contaminantes, protuberancias o vacuolas sujetas a estrés eléctrico durante periodos de tiempo prolongados, generando daños irreversibles a la aislación en el sitio en el que se presenten. Los árboles eléctricos se inician en cavidades que albergan las condiciones para la generación de descargas parciales, las que se propagan en forma tubular, conocidas como ramas. Estas ramas crecen en longitud y ancho a lo largo del tiempo hasta cruzar el aislante, generando un camino entre electrodos. De esta manera, se producirá inevitablemente la falla eléctrica del aislante y, por ende, la falla del sistema [2]. En particular, se ha observado que cables de alta tensión pueden presentar descargas parciales, las que deterioran el aislante. Este deterioro puede ser acelerado por la temperatura de funcionamiento y los transientes eléctricos. Además, es posible asociar estas descargas parciales a la existencia de árboles eléctricos [3], los cuales son indicadores del deterioro de la aislación. A través de los árboles eléctricos y las descargas parciales se puede predecir la vida útil restante de los aislantes y prevenir la falla de este.

Para estudiar este mecanismo de degradación, se ha buscado caracterizar la propagación de árboles eléctricos en distintas condiciones, relacionando el crecimiento de estos a la actividad de

descargas parciales. Para ello, se han utilizado distintas técnicas como patrones resueltos de fase (PRPD), análisis de secuencia de pulsos (PSA) series de tiempo no lineales (NL TSA) y análisis de geometrías fractales, para extraer información del estado de propagación de estos árboles. Sin embargo, a la fecha no existe un modelo analítico general capaz de modelar el fenómeno, debido principalmente a su complejidad, y naturaleza caótica y estocástica. Esto dificulta la predicción del momento en que el árbol va a cruzar la extensión del aislamiento, lo que define en la práctica el fin de la vida útil del aislante como consecuencia de la inminente ruptura dieléctrica.

Actualmente, nuevos métodos han tomado relevancia para predecir el comportamiento de una serie de fenómenos no-lineales y caóticos, donde no ha sido posible estudiarlos con métodos clásicos. Una de las herramientas que cada día cobra más fuerza son las técnicas de inteligencia artificial (IA), entre ellas, *Machine Learning*: técnicas de aprendizaje por computadora. Las técnicas de IA se han aplicado en distintas áreas del conocimiento para estudiar problemáticas tales como la predicción de velocidades de viento, disponibilidad solar y demandas energéticas, entre otras. Estas técnicas han entregado resultados satisfactorios, sin la necesidad de contar con un conocimiento pleno de la fenomenología que gobierna el fenómeno, o de modelos matemáticos que lo representen. Lo anterior es de suma utilidad debido a su carácter pragmático en la resolución de problemas, evitando tener que contar con expresiones analíticas. Con las técnicas de machine learning, se abre una nueva ventana para caracterizar el fenómeno de crecimiento de árboles eléctricos y complementarlo con los métodos ya existentes. De esta manera, con el uso de este tipo de técnicas se espera obtener una mejor descripción del fenómeno y contribuir, así, a la predicción del fin de la vida útil de los sistemas de aislamiento.

En este contexto, el objetivo de este trabajo es utilizar de técnicas de *Machine Learning* para estimar la propagación de árboles eléctricos en aislantes de resina epóxica sometidos a envejecimiento acelerado. Para ello, se cuenta con mediciones de la longitud y ancho del árbol, y del promedio de descargas parciales, obtenidos del trabajo de doctorado del Dr. Ibrahim Iddrissu (The University of Manchester),” Study of Electrical Strength and Lifetimes of Polymeric Insulation for DC Applications” [4].

El objetivo de este trabajo de memoria es aportar a la predicción de la vida útil de aislantes, a través del estudio de este fenómeno en una resina epóxica, y sentar las bases para futuras aplicaciones en el diagnóstico de equipamiento eléctrico. Esto se logrará evaluando la capacidad de predicción de técnicas de inteligencias artificial en la propagación de árboles eléctricos.

Este estudio se centrará en resina epóxica sujeta a envejecimiento acelerado por arborización eléctrica. Sin embargo, a pesar de ser un caso particular, la metodología y herramientas

desarrolladas se pueden utilizar para estudiar el envejecimiento de otros tipos de polímeros. De esta forma, se espera contribuir al área de evaluación y predicción del estado de deterioro de los aislantes, aspecto fundamental para pasar de un esquema de mantenimiento preventivo a uno basado en la condición real del equipamiento.

El objetivo general de la memoria es:

- Evaluar la capacidad de predicción de la evolución de árboles eléctricos en una resina epóxica mediante técnicas de inteligencia artificial.

Los objetivos específicos son:

- Revisar bibliografía técnica sobre la caracterización y predicción de la evolución de árboles eléctricos mediante técnicas de inteligencia artificial.
- Seleccionar características de entrada y modelos de inteligencia artificial para la predicción de la evolución de árboles eléctricos.
- Implementar modelos de inteligencia artificial apropiados, en base a la literatura técnica, para predecir la evolución de árboles eléctricos.
- Evaluar el desempeño de los modelos entrenados para la predicción de la evolución de árboles eléctricos en resina epóxica y proponer una metodología de análisis según los resultados obtenidos.

Capítulo 2

Marco teórico

En esta sección se describe el marco teórico que sustenta este trabajo. Se introducen conceptos e ideas respecto al envejecimiento de aislantes, mecanismos de envejecimiento, análisis de descargas parciales y cómo los métodos de análisis permiten el diagnóstico de los equipamientos eléctricos. Con lo anterior, se busca enmarcar el por qué el análisis de árboles eléctricos es un fenómeno que, al comprenderlo, contribuye al diagnóstico del estado de deterioro de aislantes poliméricos en equipamiento eléctricos. Además, se busca mostrar cómo las técnicas de inteligencia artificial aplicada al estudio del crecimiento de árboles eléctricos pueden aportar a la predicción del estado de propagación de estos.

2.1. Envejecimiento de aislantes

Una componente fundamental de los sistemas de transmisión y equipos eléctricos son los aislantes. La falla de un aislante puede producir importantes pérdidas económicas o humanas, por lo que es de vital importancia anticiparse y evitar este tipo de fallas. La gran mayoría de los equipos eléctricos operan durante decenas de décadas, por lo que es fundamental diagnosticar su estado, en particular los aislantes, para su reparación, reemplazo o mantención antes de que ocurra una falla.

Los factores de envejecimiento de un sistema de aislación se pueden dividir en cuatro: factores térmicos, eléctricos, mecánicos y ambientales (TEAM). Estos factores pueden actuar de manera individual o conjunta y activan distinto mecanismo de envejecimiento. Pueden producir cambios en las propiedades del material de un sistema de aislación (envejecimiento intrínseco) o producir una degradación del sistema de aislación producto de la interacción con contaminantes, defectos, protuberancias o vacíos en el material o en las interfaces de este (envejecimiento extrínseco). En general, estos defectos se introducen de forma involuntaria durante el procesado del material, la manufactura de los equipos, el transporte, la instalación o en servicio [1]. Algunos de los principales factores de envejecimiento se presentan en la Tabla 1.

Térmica	Eléctrico	Mecánico	Ambiental
-Altas temperaturas -Oscilaciones de temperatura -Bajas temperaturas	-Tensión (Alterna, continua, impulsos) -Corriente -Frecuencia	-Tensión, compresión, esfuerzos cortantes. -Fatiga, vibración.	Humedad Líquidos, Gases. Radiación

Tabla 1: Factores de envejecimiento en aislantes.

2.1.1. Descargas Parciales

De acuerdo con la norma IEC 60270, una descarga parcial se define como una “*descarga eléctrica localizada que produce la conducción parcial a través del aislante entre conductores y puede o no ocurrir adyacente a un conductor*”. Estas descargas parciales son, generalmente, consecuencias de concentraciones de estrés eléctrico en el aislante o en la superficie. Comúnmente, tales descargas aparecen como pulsos de duración menor a 1 microsegundo [5].

Las descargas parciales ocurren en cualquier sistema de aislación de equipamientos eléctricos. En general, están restringidas a una parte del material dieléctrico, que puede ser sólido, líquido o gaseoso [6]. Las descargas parciales ocurren, además, de forma aleatoria y obedecen a un fenómeno estocástico.

Las descargas parciales se pueden agrupar de acuerdo a su fuente en 3 tipos, descargas internas, superficiales y corona [7]. Las descargas internas ocurren en vacuolas o cavidades al interior de dieléctricos sólidos o líquidos. Típicamente, este tipo de descargas se originan por el desprendimiento del aislante del electrodo o por micro cavidades presentes por los procesos de manufactura y tienden a dar origen a la expansión de árboles eléctricos en la longitud de aislantes poliméricos. Las descargas superficiales aparecen en las fronteras de materiales aislantes diferentes. Son producidas por contaminación en la superficie, como partículas metálicas, produciendo caminos de baja resistencia. Finalmente, las descargas tipo corona ocurren en dieléctricos gaseosos en la presencia de campos eléctricos no homogéneos. Las figuras a continuación muestran de forma ilustrativa los tres tipos de descargas.

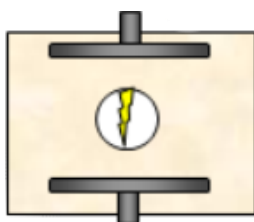


Figura 2.1: Descarga interna

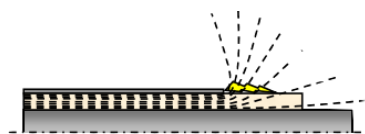


Figura 2.2: Descarga superficial

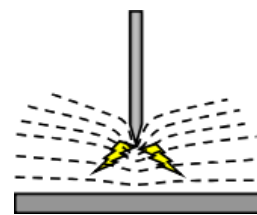


Figura 2.3: Descarga corona

2.1.2. Árboles eléctricos

Los árboles eléctricos son estructuras que crecen con formas similares a árboles botánicos. Consisten en micro canales gaseosos con o sin paredes carbonizadas que pueden ocurrir en intensificadores de estrés como protuberancias, contaminantes, vacuolas o interfaces de aislantes expuestas a estrés eléctrico prolongado [8]. En la Figura 2.4 se muestra un ejemplo de un árbol eléctrico en una resina epóxica.



Figura 2.4: Árbol eléctrico en resina epóxica [9]

Los árboles eléctricos se inician por el efecto de campos eléctricos alternos de alta divergencia, intensificados por características geométricas como una punta de electrodo o por descargas parciales en cavidades con gas. Alternativamente, pueden ser producidos por cortocircuitos o inversiones de polaridad de tensión continua [2]. El proceso de generación se puede dividir en, al menos, 2 etapas. La primera etapa, de incubación, corresponde a la etapa donde se genera el primer defecto tipo túnel, que tiene la capacidad de albergar descargas parciales. En la actualidad, aún no se sabe el proceso por el cual la inyección de carga suministra la energía al polímero para generar un daño físico en este, generando así el primer daño tipo túnel en el aislante. Algunas de las alternativas propuestas para explicar el fenómeno son la ruptura de enlaces químicos mediante radiación UV emitida durante la recombinación de carga eléctrica, la degradación química mediante impacto de electrones de alta energía cinética o la deformación mecánica del polímero. Sin embargo, lo que parece seguro es la presencia de estados de electroluminiscencia, que preceden la actividad de descargas parciales [2].

La segunda etapa es la de propagación, donde el árbol eléctrico crece a través de la aislación formando ramificaciones similares a arboles botánicos [2]. La transición desde la etapa de incubación a la de propagación de los árboles eléctricos la determina la aparición de actividad de descargas parciales, típicamente de magnitudes mayores a 0.1 pC. Estas descargas ocurren en cavidades tubulares y con el tiempo generan ramificaciones. Las descargas parciales en los canales de los árboles se pueden explicar por la teoría de ruptura dieléctrica en gases, al estar estos llenos de gas. Un modelo que explica el avance del árbol eléctrico es el modelo de descarga de avalancha [10].

El crecimiento de árboles eléctricos se puede caracterizar con modelos estocásticos o determinísticos [10]. Ejemplos de estos modelos son caracterizaciones geométricas mediante dimensión fractal [11], caracterización mediante análisis de las componentes de frecuencia de los pulsos de descargas parciales [12], análisis clásicos con diagramas resueltos en fase (PRPD),

y análisis de secuencia de pulso (PSA) [13]. La importancia de entender los árboles eléctricos es que son un indicador del estado de deterioro del aislante. En [9], se estudió la evolución de los árboles eléctricos en secciones de cables XLPE de 35 kV a través de la descomposición de la forma de onda de los pulsos de descargas parciales utilizando series de Fourier, a través de las cuales se generaron patrones tiempo-frecuencia de la forma de onda de los pulsos de descargas parciales. El objetivo del estudio fue buscar características eléctricas que explicaran la etapa de avance de los árboles eléctricos. Los resultados mostraron que las características de tiempo-frecuencia de los pulsos de descargas parciales son diferentes para cada etapa del árbol eléctrico. Otro ejemplo es el estudio publicado en [14], donde se generaron árboles eléctricos en partes de un cable XLPE de una línea de transmisión de 110 kV, a través de la inyección de descargas parciales. En el estudio, se encontró una relación entre la actividad de descargas parciales y la longitud de los árboles, la que se pudo usar para evaluar el grado de desarrollo de los árboles eléctricos y, por consiguiente, del deterioro del aislante.

2.1.3. Características relevantes

La norma IEC 60270 propone una serie de parámetros estadísticos para caracterizar las descargas parciales [5]. Los parámetros estadísticos típicamente utilizados son:

$$\text{Corriente promedio de descarga: } I = \frac{1}{T_{ref}} \sum_{i=1}^N |q_i| \quad (2.1)$$

$$\text{Potencia de descarga: } P = \frac{1}{T_{ref}} \sum_{i=1}^N q_i u_i \quad (2.2)$$

$$\text{Razón cuadrática: } D = \frac{1}{T_{ref}} \sum_{i=1}^N |q_i^2| \quad (2.3)$$

$$\text{Razón de repetición de pulso: } n = \frac{N}{T_{ref}} \quad (2.4)$$

$$\text{Magnitud de descarga media: } q_{avg} = \frac{\sum_{i=1}^N |q_i|}{N} \quad (2.5)$$

$$\text{Magnitud máxima: } q_{max} = \max(q_1, q_2, \dots, q_i) \quad (2.6)$$

Donde T_{ref} corresponde al intervalo de tiempo de análisis, q_i la magnitud aparente de la i -ésima descarga, N la cantidad de descargas registradas y u la tensión al momento de ocurrir la descarga. Además, estos parámetros se pueden caracterizar en base a su percentil de ocurrencia, por ejemplo, percentiles 25, 50, 75 y 95 de la magnitud absoluta de las descargas parciales.

Otra técnica altamente utilizada para el análisis de descargas parciales son los diagramas resueltos en fase (phase-resolved partial discharge, PRPD). Estos diagramas presentan, en un

gráfico, la actividad de descargas parciales en función de la fase de la tensión al momento de su ocurrencia. La frecuencia de la señal de referencia utilizada en el instrumento de medición para la ubicación de la fase de ocurrencia de las descargas parciales típicamente es la aplicada para la generación de las descargas parciales [15]. Lo usual para caracterizar la actividad de las descargas parciales es utilizar el valor máximo de las descargas, el promedio y la cantidad de descargas parciales. Además del diagrama, se pueden obtener distribuciones univariadas de cada parámetro:

$$\begin{aligned}
 H_{qmean}(\phi) & : \text{distribución del valor medio de la carga en cada ventana de fase} \\
 H_{qmax}(\phi) & : \text{distribución del valor máximo de carga en cada ventana de fase.} \\
 H_n(\phi) & : \text{distribución de la cantidad de descargas parciales en cada ventana de fase.}
 \end{aligned}$$

Lo usual es obtener estas distribuciones de manera diferenciada para el semiciclo positivo y el semiciclo negativo de la señal de referencia. Además, de las distribuciones se pueden obtener los momentos estadísticos, siendo los siguientes los más utilizados:

$$\text{Promedio: } \mu = \frac{\sum_{i=1}^N x_i f(x_i)}{\sum_{i=1}^N f(x_i)} \quad (2.7)$$

$$\text{Varianza: } \sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2 f(x_i)}{\sum_{i=1}^N f(x_i)} \quad (2.8)$$

$$\text{Skewness: } S_k = \frac{\sum_{i=1}^N (x_i - \mu)^3 f(x_i)}{\sigma^3 \sum_{i=1}^N f(x_i)} \quad (2.9)$$

$$\text{Kurtosis: } K_u = \frac{\sum_{i=1}^N (x_i - \mu)^4 f(x_i)}{\sigma^4 \sum_{i=1}^N f(x_i)} - 3 \quad (2.10)$$

2.1.4. Técnicas de identificación de características relevantes.

La innovación tecnológica reciente ha permitido la recolección de cantidades masivas de datos con una gran cantidad de características, lo que ha permitido el descubrimiento de nuevos conocimientos y el modelamiento de patrones mediante técnicas computacionales. Algunos de los problemas derivados de la masividad de estos datos es lo difícil que es almacenarlos y analizarlos. Por esto se hace necesario tener métodos que reduzcan la cantidad de estos con el

objetivo de reducir la información almacenada y los tiempos de cómputo en la etapa de análisis. El método que cumple esta función es la llamada *Reducción de Dimensionalidad*, la cual extrae un número reducido de características mediante la remoción de información irrelevante, redundante y/o ruidosa. Esta reducción de dimensionalidad busca disminuir la cantidad de datos a manejar a la vez de no generar una reducción significativa en la capacidad de predicción de los modelos de machine learning [16].

Por lo anterior es que, una vez obtenidas las características para representar las descargas parciales y los árboles eléctricos, es necesario identificar un subconjunto relevante. Esto debido a que, en muchos casos, existen características con una alta correlación, por lo que no aportan información nueva. Por otra parte, utilizar demasiadas características puede llevar a problemas de “sobreajuste”, disminuyendo la capacidad de generalización de los algoritmos.

Existen básicamente dos maneras para extraer características relevantes: técnicas de “selección” y técnicas de “extracción”. Por una parte, las técnicas de selección buscan encontrar un subconjunto reducido de características que contengan la mayor cantidad de información posible, descartando el resto. Esta selección de características del conjunto original permite una exploración intuitiva del problema a través de los datos, pues al conocer las características originales relevantes estas se pueden asociar al fenómeno físico en análisis logrando un mayor entendimiento del problema [17]. Algunos ejemplos de esta técnica es la denominada “Selección Recursiva de Características” la cual selecciona características de forma secuencial basada en métricas de importancia para generar un ranking indicando cuales características podrían ser relevantes. Por su parte, las técnicas de extracción buscan extraer o derivar características a partir del conjunto original, generando, entonces, un nuevo conjunto de características. Un ejemplo de esta técnica es el “Análisis de Componentes Principales” [18].

2.2. Técnicas de inteligencia artificial

Las técnicas de inteligencia artificial son algoritmos útiles para identificar patrones, predecir o clasificar un fenómeno, en base a datos, en particular cuando los fenómenos en estudio no están descritos por un modelo analítico (o el modelo analítico es de alta complejidad). La ventaja es que, para estudiar el fenómeno, se generan modelos de tipo “caja negra” que estiman una salida, en base a un conjunto de datos de entrada, por lo que no es necesario tener un modelo analítico.

2.2.1. Técnicas de aprendizaje

Los métodos de inteligencia artificial se pueden dividir en 3 grandes grupos en función de cómo “aprenden” de los datos. Estos métodos son “aprendizaje supervisado”, “aprendizaje no supervisado” y “aprendizaje reforzado”.

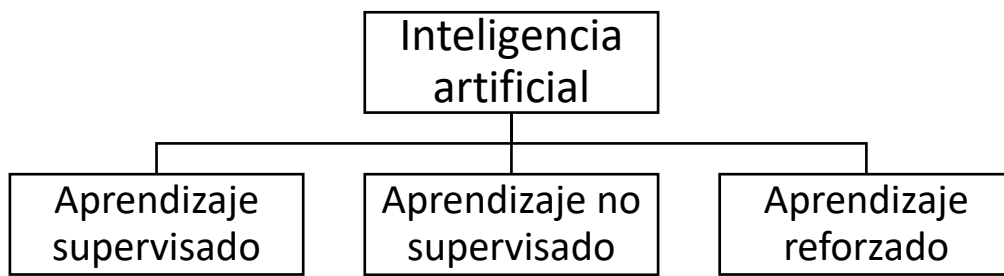


Figura 2.5: Grupos inteligencia artificial

2.2.1.1 Aprendizaje supervisado

El aprendizaje supervisado es una técnica para deducir una función (es decir, el modelo) a partir de datos de entrenamiento, los cuales consisten en datos de entrada y salida. La función a deducir consiste en un mapeo entre los datos de entrada y los de salida. Los datos de salida pueden ser variables continuas, en cuyo caso se habla de un problema de “regresión”, o una etiqueta de clases, en cuyo caso se habla de problemas de “clasificación”. Por ejemplo, para el caso de la evolución de árboles eléctricos, los datos de entrada podrían ser la cantidad de descargas parciales por minuto y los datos de salida podrían ser la longitud del árbol eléctrico (problema de regresión) y/o la etapa de evolución del árbol eléctrico (para un problema de clasificación), entre otros. Una vez obtenido el modelo, se puede estimar el valor de la(s) salida(s) a partir de nuevos datos de entrada. El término supervisado hace referencia a que los datos de salida son conocidos por la inteligencia artificial al momento de entrenar el modelo, es decir, el modelo se guía por estos valores para generar una predicción.

2.2.1.2 Aprendizaje no supervisado

El aprendizaje no supervisado se caracteriza por tratar con datos no etiquetados o datos con estructuras desconocidas. En este caso, el objetivo es extraer información relevante de los datos, como por ejemplo identificar patrones a partir de similitudes y diferencias entre las características de los datos. Una de las técnicas de inteligencia artificial más utilizada en este tipo de aprendizaje es el “*clustering*”, que busca agrupar los datos en grupos (o “*clusters*”) con características similares entre sí. De este modo, un modelo puede predecir la característica o *cluster* al que pertenece un nuevo grupo de datos a partir de las características o *clusters* detectados en los datos de entrenamiento.

2.2.1.3 Aprendizaje reforzado

El aprendizaje reforzado es una técnica que busca determinar qué “*acciones*” debe realizar un “*agente*” dentro de un “*ambiente*” dinámico con tal de maximizar una recompensa. Un ejemplo de este tipo de aprendizaje es el de un motor de ajedrez. Aquí el sistema decide qué serie de movimientos (*acciones*) debe realizar un jugador (*agente*) dependiendo del estado del tablero.

(ambiente), y la recompensa se puede definir como ganar o perder al final del juego, o la calidad de una jugada en particular.

2.2.2. Tipos de modelos

Existe una enorme cantidad de modelos de inteligencia artificial, por ejemplo, Árboles de Decisión, Random Forest, Support Vector Machine, K-Nearest-Neighbor, Regresión Logística (solo por nombrar algunos). A continuación, se describirán los modelos de tipo Random Forest y Redes Neuronales (ya que son los que se usaron en este trabajo) y Árboles de Decisión (ya que es el modelo básico en el que se basa un Random Forest).

2.2.2.1 Árboles de decisión

Un Árbol de decisión es un modelo de machine learning que utiliza aprendizaje supervisado. Este método se basa en separar los datos mediante simples decisiones lógicas a partir de las características de las muestras, para así separarlas en sus respectivas clases (para clasificación) o rangos de valores (para regresión).

Una manera simple de entender cómo funciona un árbol de decisión es tomar el ejemplo de una clasificación binaria a partir de diferentes características de entrada. En principio, el algoritmo busca el límite que mejor separa los datos en ambas clases (en función de la característica de entradas que mejor logra esta división). Para esto, el algoritmo busca reducir la “impureza” de los grupos creados tras la división. La “impureza” mide qué tan mal se clasificaron los datos entre las dos clases. Mientras más datos de una clase quedaron en el grupo de la otra clase, más “impuro” es ese grupo. La decisión que genera los grupos resultantes define un “nodo”, mientras que los grupos resultantes (y sus subsecuentes decisiones) definen las “ramas”. Hecha una división, se repite el mismo proceso sobre los grupos resultantes, es decir, se busca otro límite (usando cualquiera de las características de los datos) para separar los datos en las dos clases. Notar que este límite puede ser en cualquier característica de los datos, no necesariamente en la característica utilizada en el nodo anterior. Estas nuevas decisiones generan nuevos nodos con sus respectivas ramas. Este proceso idealmente se repite hasta lograr separar los datos en grupos “puros” (es decir, grupos que contienen elementos de una sola clase).

Para problemas multiclase este método funciona de igual forma. La única diferencia es que en cada uno de los nodos hijos habrá más de 2 clases para el grupo de muestras.

Las medidas de impureza más comúnmente usadas en árboles de decisión de clasificación son el “Índice de Gini”, la “entropía”, y el “error de clasificación”. En esta investigación se usó como

medida de impureza el índice de Gini. La formulación matemática de este es la que se muestra a continuación:

$$I_G = 1 - \sum_{i=1}^c p(i|t)^2 \quad (2.11)$$

Donde c es el número de clases y $p(i|t)$ es la proporción de las muestras que pertenecen a la clase c para un nodo particular t . Si las muestras en el nodo pertenecen solo a una clase (digamos la clase C_j) entonces se tiene una proporción p_j igual a uno y el resto de los términos $p_{i \neq j}$ se anulan, dando un índice de Gini igual a cero. Evidentemente se busca que I_G sea nulo, es decir, que el nodo sea “puro”.

2.2.2.2 Random forest

En este trabajo se utilizó el método de “Bosque Aleatorio” (*Random Forest*), el cual es un modelo “ensemble” de varios Árboles de Decisión (es decir, el resultado se compone de la composición de varios resultados provenientes de Árboles de Decisión diferentes). Los diferentes Árboles de Decisión que componen el bosque aleatorio se generan a partir de diferentes subgrupos del set de datos total de entrenamiento. Este algoritmo se puede esquematizar de la siguiente manera:

1. Se seleccionan aleatoriamente n muestras (con reemplazo) del set de datos de entrenamiento.
2. Se crea un árbol de decisión a partir de las n muestras. En cada nodo:
 - a. Se seleccionan d características sin reemplazo
 - b. Se divide el nodo usando la característica que proporcione la mejor división minimizando la impureza.
3. Se repiten los pasos 1 y 2 k de veces, generando un número k de Árboles de Decisión.
4. La predicción final del Random Forest combina la predicción de cada Árbol de Decisión mediante voto mayoritario para asignar la clase (es decir, a cada muestra le asigna la clase “votada” más veces por cada Árbol).

2.2.2.3 Redes Neuronales

Las Redes Neuronales son estructuras matemáticas típicamente compuestas por *capas*, las que a su vez se componen por *neuronas* (también llamadas *nodos*). Cada nodo realiza una suma ponderada de las características de entrada x_i , asignándole a cada característica un *peso* w_i , para luego sumar un último término, comúnmente llamado *sesgo* b (o *bias* en inglés), tal como se muestra a continuación.

$$\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

$$z = \mathbf{w}^T \mathbf{x} = w_1 x_1 + \cdots + w_m x_m + b$$

Para terminar, el resultado anterior z se entrega a función de “activación” (que suele colapsar la suma a algún resultado binario, por ejemplo 1 y 0, para reflejar si la neurona se “activó” o no). Las funciones de activación tradicionales son la función sigmoide y la tangente hiperbólica, sin embargo, en el último tiempo ha ganado popularidad la función *ReLU* (indicada a continuación) debido a su desempeño muchas veces superior a otras activaciones.

$$ReLU(z) = \begin{cases} z, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

Las redes neuronales se componen como mínimo de una *capa de entrada* (que simplemente recibe las características de los datos de entrada, sin pesar los datos y sin función de activación, por lo que muchas veces no se considera como una capa de la red neuronal propiamente tal) y una *capa de salida* (donde sí se pesan los datos y sí hay una activación). La versión más simple de una red neuronal mínima es en la que la capa de salida contiene un solo nodo. Si la función de activación es binaria (resultando, por ejemplo, en 0 o 1, dependiendo del valor de z) entonces a este sistema se la llama *perceptrón*. Por otro lado, si la activación es lineal (equivalente a no aplicar una función de activación) entonces la red neuronal es equivalente a una regresión lineal.

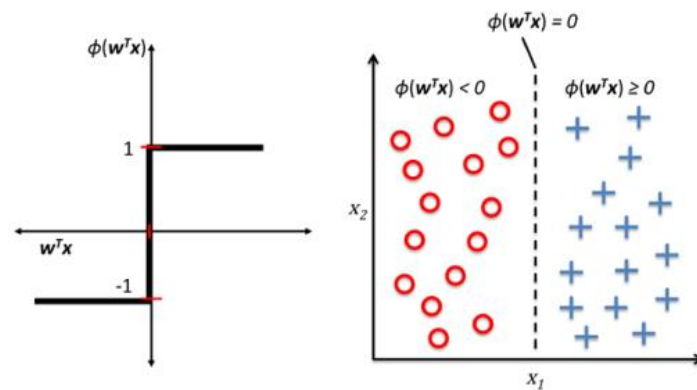


Figura 2.6: Esquema de respuesta de un perceptrón ([19])

Un perceptrón es un modelo de escaso interés práctico debido a su baja capacidad de ajuste a datos complejos. La potencia de las redes neuronales proviene de utilizar varias capas con varias neuronas, con lo que logra modelar complejas relaciones no lineales. A las capas internas (ubicadas entre la capa de entrada y la de salida) se las suele llamar *capas ocultas* (*hidden layers*). Lo usual es que todas las neuronas de una capa reciban la salida de todas las neuronas

de la capa anterior. A estas capas se las llama *densas* o *completamente conectadas* (*dense* o *fully connected*, en inglés). En el caso más simple (el de las redes *feed-forward*) los datos de salida de una capa son la entrada de la capa siguiente. Otro tipo popular son las *redes recurrentes*, donde los nodos, (a diferencia de en las *feed-forward*) reciben información tanto de los nodos anteriores como de ellos mismos.

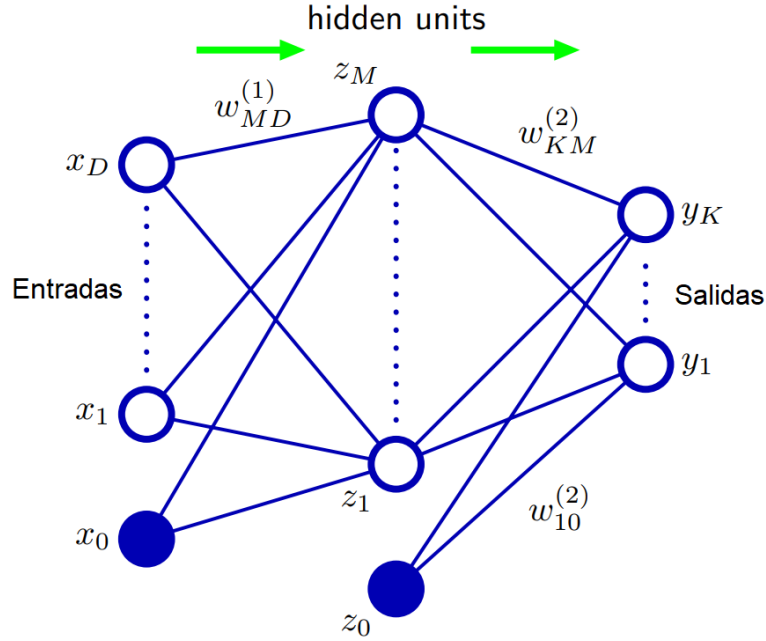


Figura 2.7: Esquema red neuronal densa feed-forward de una capa oculta ([20])

Los modelos de redes neuronales usados en este trabajo se implementaron mediante la librería de Python Tensorflow [21]. Estos modelos son de tipo feed-forward de capas densas (como el de la Figura 2.7) con activación ReLU para las capas ocultas y activación "softmax" para la capa de salida. Se utilizó la función softmax ya que es específica para los problemas de clasificación multiclase [20]. La función softmax está definida por la siguiente formula:

$$\text{softmax}(z_j) = \frac{e^{z_j}}{\sum_{i=1}^C e^{z_i}} \quad (2.12)$$

Donde $i = \{1, \dots, C\}$, C es el número de clases y z_i es la combinación lineal de la salida de la red para la clase i . El resultado para cada clase está en el rango $[0, 1]$ y la suma de todas ellas siempre es igual a 1 (es decir, está normalizado). Esto permite interpretar a $\text{softmax}(z_i)$ como la probabilidad de pertenecer a la clase i . De esta manera, cada dato se asoció a la clase correspondiente al mayor valor de $\text{softmax}(z)$ predicho.

Las redes neuronales se entrenaron mediante *gradiente descendiente* con el método de *retropropagación del error* [19]. Estos métodos funcionan de manera iterativa. Es decir, al

entrenar una red neuronal se define un número de “épocas”. En cada época se actualizan los pesos y los sesgos de las neuronas con tal de disminuir una *función de pérdida* (usualmente alguna medida del error). En las redes neuronales utilizadas en este trabajo se utilizó la *entropía cruzada* como pérdida (apropiada para ser utilizada junto a softmax en los problemas de clasificación) y el método ADAM (un tipo de gradiente descendiente popular por su efectividad) para ajustar el modelo.

2.3. Estado del arte de técnicas de inteligencia artificial aplicada a predicción de envejecimiento de aislantes

Como se indicó en la sección 2.1.2, en la actualidad no existen métodos analíticos que permitan establecer relaciones entre el estado de deterioro en aislantes poliméricos y las descargas parciales. Por su parte, el estado de deterioro se puede inferir a través de la propagación de los árboles eléctricos. Poder establecer una relación de este tipo permitiría realizar acciones preventivas, antes de que se produzca una ruptura del material aislante. Uno de estos acercamientos son las técnicas de inteligencia artificial, las que han entregado buenos resultados para problemas de naturaleza estocástica, como lo son las descargas parciales y los árboles eléctricos.

La literatura técnica en cuanto al uso de técnicas de inteligencia artificial para la aplicación en ingeniería eléctrica es vasta. Se han utilizado, por ejemplo, redes neuronales y arboles de decisión para la predicción de demandas, perfiles de generación solar, predicción de la vida de equipamientos eléctricos, perfiles de viento y otros [22] [23] [24] [25]. Sin embargo, los estudios de árboles eléctricos utilizando técnicas de inteligencia artificial son escasos. Resulta atractivo, entonces, explorar técnicas de inteligencia artificial para predecir la evolución de árboles eléctricos y, así, la vida útil de equipamientos eléctricos que utilicen aislantes poliméricos.

En cuanto a la predicción de la vida útil de equipos eléctricos, en [26] se presenta una propuesta para la predicción de la vida de cables de polietileno reticulado (XLPE). En este trabajo, se clasificaron 2500 muestras de cable XLPE de 15 kV, utilizando un índice de salud. Este índice fue generado a partir de datos de las descargas parciales, tiempo de servicio del cable, datos visuales, corrosión y otros. Para la predicción, se entrenaron 4 modelos de clasificación multiclase de 5 etiquetas: Support vector machine, KNN (K- Nearest neighbor), red neuronal artificial y naive bayes. El entrenamiento se realizó dividiendo las muestras en una razón 8:2 entre entrenamiento y prueba y se utilizó un esquema de validación cruzada K-fold de 10 pliegues con métricas de exactitud, precisión, recall y F1 Score. Los resultados obtenidos se

muestran en la Figura 2.8. Como se puede ver de esta figura, los resultados mostraron una alta capacidad de predicción en clasificar el índice de salud de las muestras de cable XLPE.

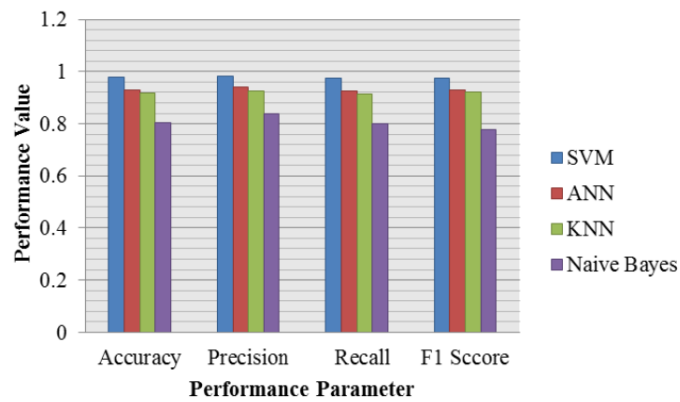


Figura 2.8: Resultado predicción de estado de salud de cables XLPE (fuente [26])

Otro ejemplo se presenta [27], para la predicción de la vida útil de aislantes en motores sumergibles. Los motores sumergibles están sujetos a distintos factores de envejecimiento que producen debilitamiento en los aislantes. Algunos de estos factores son la temperatura, tensión, y presión del agua. Para el estudio, se generaron datos de motores sumergibles de 6kV, mediante envejecimiento acelerado. Se utilizó una red neuronal con retropropagación, cuyos datos de entrada fueron la temperatura, tensión y presión de agua. El objetivo de la red neuronal fue estimar la vida útil del motor a partir de los datos de entrenamiento. Se obtuvo que el modelo tuvo errores de predicción entre -0.2 a 0.2 años. Ejemplo de las características operacionales a las cuales fueron sometidos los motores se muestran en la Tabla 2

Input			Output
Tensión kV	Temperatura °C	Presión de agua (MPa)	Vida (año)
13.30	80.50	6.25	20.15
13.25	80.43	6.23	20.21
13.32	80.43	6.23	19.89
13.28	80.53	6.24	20.13
13.40	80.52	6.25	20.22
13.37	80.42	6.21	20.12

Tabla 2: Ejemplo de datos de motores sumergibles utilizados en [27]

En [28], se estudió el progreso de árboles eléctricos en un aislante sólido de polietileno reticulado (XLPE). Para ello, se crearon 3 tipos de muestras con defectos artificiales, tal como se muestra en la Figura 2.9. El primer modelo consiste en una aguja embebida sin defecto artificial; el segundo modelo consiste en una aguja embebida con un vacío en la punta. Finalmente, el tercer modelo consiste en una aguja embebida con una partícula metálica. En el desarrollo experimental, se registró la magnitud, fase y forma de onda de las descargas parciales durante la expansión del árbol. Estas señales se caracterizaron utilizando dos distribuciones bidimensionales de los diagramas resueltos en fase, $\varphi - n$ (fase – cantidad de descargas) y $\varphi - q_{max}$ (fase – magnitud de descarga máxima). Con los datos obtenidos, se implementó un modelo de inteligencia artificial correspondiente a una red neuronal de clasificación, para asociar la fuente de descargas parciales a uno de los 3 tipos de defectos simulados en las muestras. Los datos de entrada corresponden a las características extraídas de las mediciones de descargas parciales desde el 70% al 90% de la expansión del árbol eléctrico. El objetivo de la red neuronal fue identificar, a través de los datos, a qué tipo de defecto corresponde la señal. Como resultado, se obtuvo una exactitud del 99% en la identificación de la fuente del árbol eléctrico.

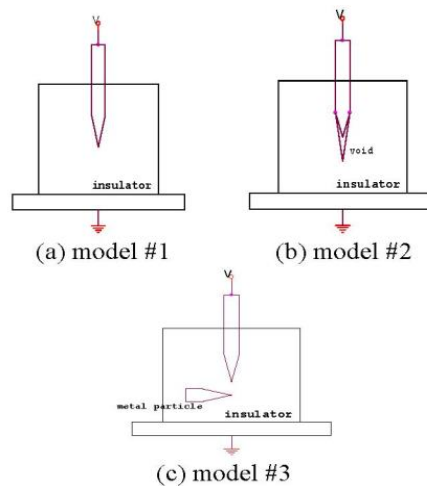


Figura 2.9: Modelos de falla para árbol eléctrico utilizado en [28]

Otra investigación de mayor relevancia se puede ver en [29]. En esta investigación se usaron redes neuronales de clasificación para identificar la etapa de evolución de árboles eléctricos en polietileno, mediante señales de descargas parciales medidas con una antena de ultra alta frecuencia (ultra-high frequency, UHF) con una detección de señales máxima de 4GHz. Para ello, se registraron mediciones de los pulsos de las descargas parciales durante el avance del árbol eléctrico, aplicando descomposiciones de wavelet a cada señal. Posteriormente, el árbol se segmentó en 4 etapas en base a su crecimiento: etapa inicial, etapa de crecimiento, etapa de

crecimiento retardado y etapa de crecimiento rápido. Las variables de entrada para la clasificación fueron parámetros estadísticos de las distribuciones de energía de la descomposición de wavelet, utilizándose en total 28 características por pulso. El objetivo de predicción fue la etapa de crecimiento del árbol eléctrico. Lo anterior se implementó mediante la técnica BPNN (back propagation neural network) de 3 capas, obteniéndose resultados de predicción con una exactitud cercana al 99%.

Los trabajos expuestos en esta sección muestran que las técnicas de inteligencia artificial podrían dar buenos resultados en predecir la vida útil de equipamientos eléctricos y, en específico, para el diagnóstico de árboles eléctricos. En especial, en [29] se lograron resultados prometedores en cuanto a la identificación de la etapa de crecimiento para árboles eléctricos, mediante el uso de variables eléctricas, lo que es el foco de esta investigación.

Los documentos revisados, a pesar de ser diversos y no enfocarse únicamente a los árboles eléctricos, coinciden en ciertos puntos. Se vio que, para la predicción de la vida de equipamientos y árboles eléctricos, se han utilizado técnicas de inteligencia artificial de clasificación multiclase, asignando estados (etiquetas) al nivel de deterioro, es decir, por ejemplo, estado bueno, malo y muy malo. Además, se repite el uso de redes neuronales, debido a la flexibilidad con la que se pueden ajustar a distintos fenómenos.

Capítulo 3

Procedimiento experimental

En este capítulo se describen los datos utilizados en esta memoria y los análisis realizados para la implementación de los modelos de inteligencia artificial. Además, se muestran aspectos relacionados con el manejo de datos y las características extraídas.

3.1. Probetas de ensayo

Los datos utilizados en esta memoria fueron obtenidos de una sección del estudio del Dr. Ibrahim Iddrissu [4] sobre el tiempo de vida de aislantes poliméricos para aplicaciones de corriente continua. Dentro de su estudio, investigó la propagación de árboles eléctricos en resina epóxica, tomando datos de las descargas parciales, ancho y longitud de los árboles en el tiempo.

El experimento realizado por el Dr. Ibrahim tuvo dos etapas. Una etapa de iniciación rápida de árboles, y una etapa de propagación monitoreada por detección óptica, y de medición de descargas parciales. La etapa de iniciación tiene como fin generar un árbol en la probeta y la segunda tiene como objetivo propagar los árboles a partir del defecto inicial generado en la etapa de incubación. Este trabajo se enfocó solo en la etapa de propagación.

Las muestras utilizadas para la incubación y propagación de los árboles corresponden a cubos de resina epóxica de 25 milímetros de lado con una aguja embebida fabricada por la empresa Ogura [30], con la cara del plano de tierra metalizada con aluminio para un buen contacto con el electrodo de tierra. La longitud por la que se puede propagar el árbol es el espacio entre la punta de la aguja y el plano de tierra, teniendo un espacio de alrededor de 2 milímetros entre ambos. Un esquema tipo de las probetas se muestra en la Figura 3.1.

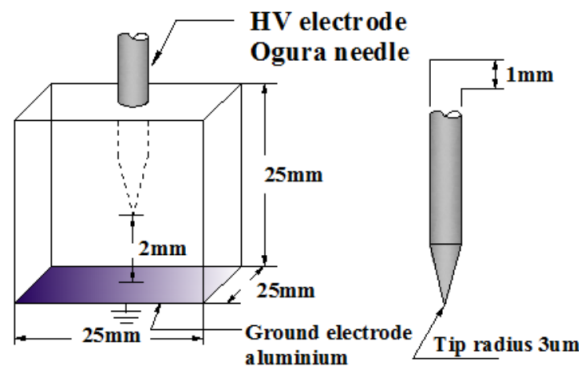


Figura 3.1: Esquema muestra (fuente [4])

Cabe destacar que el largo final de los árboles, al momento de tocar con la cara de la probeta, medido a través de los registros de la cámara, no siempre fue de 2 milímetros. Esto se debe por variaciones en la fabricación de las probetas, que hacen que la distancia entre la aguja y el plano no siempre sea de 2 milímetros, o por el paralaje en la toma de imágenes, que no siempre es el mismo.

3.2. Ruido de medición

El ruido de medición siempre es un problema al montar experimentos que involucren la detección de descargas parciales. Esto se debe a que el orden de magnitud de las corrientes generadas por una descarga es similar al ruido producido por el equipamiento eléctrico y las ondas electromagnéticas presentes en el sitio donde se está realizando el experimento. Para eliminar el ruido en la obtención de los datos, se empleó un circuito de medición balanceado acorde a la norma IEC 60270 [5]. El montaje experimental logró captar ruido entre 0.25 pC a 0.35 pC, por lo que se decidió considerar solo descargas parciales con magnitudes de 0.4 pC o superiores, como se indica en [4].

3.3. Datos disponibles

En total, se realizaron tres tipos de experimentos, cada uno con distinta tensión de entrada, aplicada en la aguja embebida:

- tensión AC pura de 15kV peak
- tensión AC de 15kV peak con offset DC de +15kV
- Tensión AC de 15 kV peak con offset DC de -15kV

En total, se obtuvieron 10 probeta para el caso de inyección de una tensión AC pura, 9 para una tensión AC con offset DC de +15kV y 7 para una tensión AC con offset DC de -15kV. Es decir, en total se obtuvieron 26 muestras de árboles eléctricos.

En cada ensayo, se registraron las magnitudes de las descargas parciales, y el largo y ancho de los árboles eléctricos cada 5 minutos. Con estos datos, se obtuvo también el tiempo que pasó desde el inicio del árbol, hasta el momento en donde se produjo la ruptura dieléctrica. Cabe destacar que no todos los ensayos terminaron en una ruptura dieléctrica, por lo que en dichos casos el tiempo final registrado corresponde al término del ensayo, del orden de las horas. Debido a esto, en este trabajo se registró también el tiempo transcurrido hasta la llegada del árbol al contraelectrodo. La ventaja es que, una vez que el árbol llega al contraelectrodo, el tiempo que demora en producirse la ruptura dieléctrica es aleatorio y no muestra una relación causal con la actividad de descargas parciales. Por otro lado, la llegada del árbol al contraelectrodo marca en términos prácticos el fin de la vida útil de un aislante.

Capítulo 4

Modelos y Resultados de predicción

En este capítulo se presentan los modelos utilizados para evaluar la capacidad de predicción de la evolución de árboles eléctricos en resina epóxica mediante técnicas de inteligencia artificial y los resultados de predicción obtenidos. Se describen los modelos utilizados, así como la metodología de análisis para evaluar la capacidad de predicción de árboles eléctricos. Además, se describen métodos para identificar las características importantes para los modelos escogidos.

4.1. Modelos propuestos

La capacidad de predicción de la evolución de los árboles eléctricos se evaluó mediante tres modelos de machine learning.

1. Modelo random forest: Random forest mediante la librería scikit learn con 100 árboles de decisión; criterio de la impureza de Gini para la división en cada rama de los árboles; el número de características a considerar para buscar la mejor división es la raíz cuadrada del número de características totales; el submuestreo para los datos de entrenamiento de cada árbol se realizó mediante “Bootstrap”. El resto de los parámetros utilizados son los fijados por defecto en la librería scikit para random forest [31].
2. Modelo 1: Red neuronal de 3 capas, con una capa oculta, la que cuenta con 30 unidades. Para las unidades de la capa oculta se utilizó la función de activación ReLU, y para las variables de salida se utilizó la función de activación softmax.

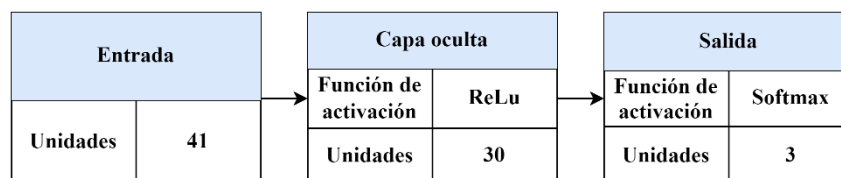


Figura 4.1: Esquema red neuronal modelo 1 generado por TensorFlow

3. Modelo 2: Red neuronal de 4 capas con 2 capas ocultas, la que cuenta con 30 unidades en la primera capa, y 15 unidades en la segunda. Para las unidades de las capas ocultas se utilizó la función de activación ReLU, y para las variables de salida se utilizó la función de activación softmax.

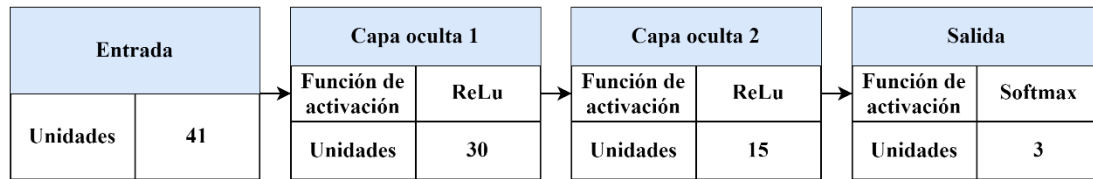


Figura 4.2: Esquema red neuronal modelo 2 generado por TensorFlow

Los datos de entrada a los modelos de red neuronal son los valores estandarizados de las características que se quieren utilizar para la predicción. En este contexto, los datos de cada característica se centraron, para que tuvieran una media 0 y una desviación estándar 1. La ecuación que representa este procedimiento es la siguiente:

$$x_{std} = \frac{x - \mu_x}{\sigma_x} \quad (4.1)$$

Donde x es el valor de la característica, μ_x es la media y σ_x la desviación estándar. El preprocesado de los datos mediante estandarización es necesario debido a que en el perceptrón se realiza una combinación lineal de las características mediante los pesos del modelo. Entonces, si las características tienen valores en ordenes de magnitud diferentes entre sí, al inicializar los pesos en la etapa de entrenamiento (típicamente todos en valor 1), las características con órdenes de magnitud mayores tendrán una dominancia intrínseca respecto a las otras. Esto puede generar que el modelo sea sensible solo a estas características. Asimismo, tener características con distintos órdenes de magnitud puede introducir problemas numéricos.

Para identificar la calidad de la información contenida en las características IEC y PRPD, desde el punto de vista de capacidad de predicción, se entrenaron tres modelos. El primer modelo utiliza como datos de entrada sólo las características IEC, el segundo sólo las características PRPD y el tercero utiliza ambos tipos de características. Para determinar la capacidad de predicción de cada modelo, se utilizó la tasa de verdaderos positivos, es decir, la cantidad de veces que el modelo predijo correctamente la salida sobre el total de las predicciones. El procedimiento específico para entrenar cada modelo fue el siguiente:

- Modelos de Random Forest: se realizó una validación cruzada *k-fold* de 5 pliegues, obteniendo los valores promedio de cada caso. Se utilizó una relación de 80% y 20% para los datos de entrenamiento y prueba, respectivamente.
- Redes neuronales: Se entrenaron los modelos a través 1000 épocas. Se guardó el modelo que minimiza la función de costos en los datos de validación. La métrica utilizada para el entrenamiento de retropropagación fue *Sparse categorical cross entropy*, una medida de la entropía cruzada para modelos de clasificación multiclase.

Luego se evaluó la capacidad de predicción en los datos de prueba y se obtuvieron las matrices de confusión para cada caso.

Posterior a las etapas de entrenamiento y evaluación, se realizó un análisis de importancia de características mediante la técnica de Gini para Random Forest y la importancia de permutación para redes neuronales. Una vez definidos los puntajes de importancia, se realizó una selección de características mediante eliminación recursiva de características (*Recursive Feature Elimination, RFE*), finalizando con una optimización de hiperparámetros mediante una búsqueda de rejilla para el modelo 2 de redes neuronal. La búsqueda de rejilla se realizó mediante una búsqueda para 1, 2, y 3 capas ocultas con 10, 20, 30 y 40 unidades ocultas en cada una.

4.1.1. Características relevantes

La selección de características es un método para encontrar qué características son relevantes para el modelo de inteligencia artificial. Es importante hacer notar que esta importancia está condicionada con la capacidad de predicción del modelo utilizado para identificar estas características. Es decir, si el modelo tiene una baja capacidad de predicción, entonces las características identificadas como relevantes no necesariamente lo serán. Para la identificación de las características relevantes se usaron métodos independientes para Random Forest y redes neuronales. Tal como se explicó en la sección 2.2.2.2, el modelo de Random Forest obtiene en cada división de sus árboles de decisión, una ganancia de información medida por la métrica de impureza. En este caso particular dicha métrica corresponde al índice de Gini. Para obtener un índice de importancia se usa, entonces, la reducción promedio de impureza, que corresponde al promedio de la diferencia entre la impureza del nodo padre y la suma ponderada por la proporción de muestras de la impureza de los nodos hijos. Para las redes neuronales, se utilizó una técnica de permutación para identificar las características relevantes. Esta métrica consiste en evaluar el desempeño del modelo al intercambiar, aleatoriamente, los datos de la característica de análisis entre todas las muestras del set de prueba, para así romper la relación creada durante el entrenamiento entre la característica a analizar y el modelo de inteligencia artificial. El puntaje obtenido corresponde a la diferencia entre la métrica de evaluación con los datos de la característica sin intercambiar, y la métrica de evaluación con los datos de la característica intercambiados aleatoriamente entre todas las muestras del set de prueba. Así una disminución en la métrica de evaluación indica que el modelo es dependiente de la característica en análisis para realizar una correcta predicción.

4.1.2. Selección de características

Se realizó una eliminación recursiva de características (Recursive feature elimination, RFE) para cada modelo. Este método se basa en entrenar y validar un modelo de forma iterativa obteniendo en cada iteración la importancia de características para, así, en la siguiente iteración eliminar la característica menos relevante al momento de entrenar y evaluar el modelo. Este proceso iterativo se repite con todas las características hasta quedar con solo una. Este método se realiza con la intención de obtener un modelo con un set de datos con características reducidas y una mejora en la capacidad de predicción. Para la eliminación recursiva de características se utilizó el modelo de random forest de 100 árboles y para las redes neuronales se utilizó el modelo 2 de redes neuronales, correspondiente a la red neuronal de 2 capas ocultas.

4.1.3. Etiquetas de estado

Las etapas de evolución de los árboles eléctricos se pueden dividir en 3; la etapa de incubación (inception stage) donde se genera el primer daño de tipo cavidad tubular en el aislante con tiempos de incubación de alta variabilidad entre muestras; la etapa de propagación (propagation stage) caracterizada por actividad de descargas parciales y un crecimiento del árbol eléctrico y; la etapa de fuga (runaway stage), donde, aumenta la actividad de descargas parciales. En esta última etapa, la velocidad de expansión del árbol eléctrico aumenta y el árbol eléctrico cruza la región del aislante hacia el contraelectrodo. Estas etapas se resumen en el esquema de la Figura 4.3 de longitud del árbol versus tiempo.

Las etapas definidas en este trabajo se realizaron basándose a modo de ejemplo en el comportamiento de las descargas parciales en el tiempo presentado en la Figura 4.4. Siendo las etapas las siguientes:

1. Etapa 1: caracterizada por alta actividad de descargas parciales con elevada magnitud y un crecimiento rápido del árbol eléctrico. Corresponde a la etapa 2 de la Figura 4.4.
2. Etapa 2: se observa baja actividad de descargas parciales con magnitudes pequeñas y un aumento de la velocidad de propagación del árbol eléctrico. Esta etapa corresponde a la etapa 3 de la Figura 4.4.
3. Etapa 3: Aumento de la actividad de descargas parciales con un aumento de las magnitudes de las descargas parciales y disminución de la velocidad de propagación del árbol eléctrico. Corresponde a la etapa 4 de la Figura 4.4.

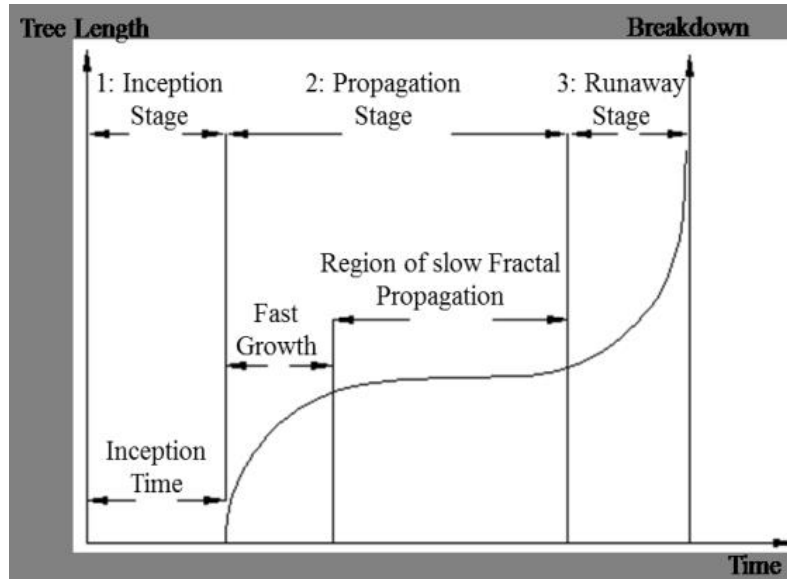


Figura 4.3: Etapas árboles eléctricos usada en [4]

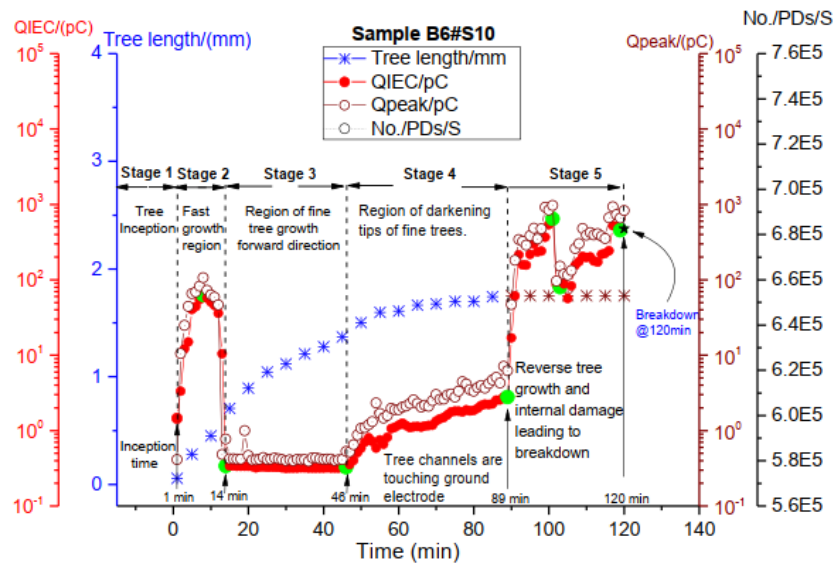


Figura 4.4: Etapas propagación árbol eléctrico usada en [4]

En este trabajo, no se consideró ni la etapa 1 ni la etapa 5. Por una parte, la etapa 1 corresponde al periodo de incubación del árbol eléctrico, el cual está fuera del análisis de esta investigación. Por otra parte, la etapa 5 es posterior a la llegada del árbol eléctrico al contra electrodo, por lo que el material ya perdió su vida útil.

La determinación de la etapa del estado de los árboles (etiqueta) se hizo de acuerdo con las etapas de la Figura 4.4, de manera manual. Por cada muestra, se determinó el instante de tiempo cuando el árbol pasó de la etapa 1 a la etapa 2 y el instante de tiempo cuando el árbol pasó de

etapa 2 a la etapa 3. De esta forma, por cada árbol, los datos registrados durante su evolución se etiquetaron de acuerdo a la etapa de crecimiento del árbol.

4.2. Resultados random forest

La tabla 3 muestra los resultados de predicción con random forest, cuando se utilizan solo las características definidas por la IEC, solo las características extraídas de PRPD y ambos tipos de características. Como se puede ver, el mejor modelo fue el que utilizó ambos tipos de características, obteniendo una exactitud en la predicción de 90%. Esto demuestra que la información contenida en las características IEC y PRPD se complementan entre sí.

Al observar las matrices de confusión de la Figura 4.5 y Figura 4.6, se observa que los modelos entrenados con características IEC e PRPD no tuvieron mayores diferencias en su capacidad de predecir el estado del árbol. En ambos casos, la exactitud en la predicción varió entre un 87,67% y 91,25% para el modelo que utiliza las características IEC, y entre 87,67% y 95% para el modelo que utiliza las características PRPD.

Observando la matriz de confusión de la Figura 4.7 para el caso IEC + PRPD se tiene una mayor exactitud en la clasificación de cada clase respecto a los casos IEC y PRPD, variando entre 94.52% y 97.53%.

La Figura 4.8 muestra la importancia de las características, de acuerdo al modelo de random forest. Como se puede ver, las 5 características más importantes fueron PD_mean_pc, PD_max_pc, pos_mean_area, pos_n_area y neg_max_kurt. Lo relevante es que las primeras 2 características son de tipo IEC y las últimas 3 se obtuvieron del análisis PRPD. Además, hay 2 características de área que no pertenecen al análisis típico de los diagramas resueltos en fase, lo que significa que todavía es posible extraer información mediante características no convencionales con las herramientas actuales para el análisis de descargas parciales.

Modelo	Tipo de característica		
	IEC	PRPD	Todas
Random Forest	0.8943	0.9136	0.9191

Tabla 3: Evaluación exactitud modelo random forest para los tipos de características.

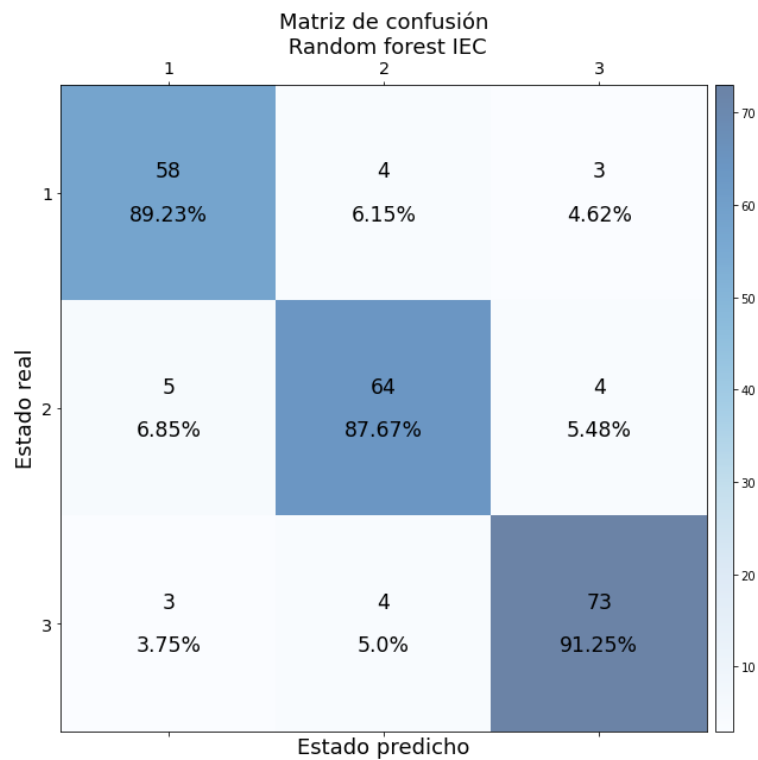


Figura 4.5: Matriz de confusión para random forest para características IEC

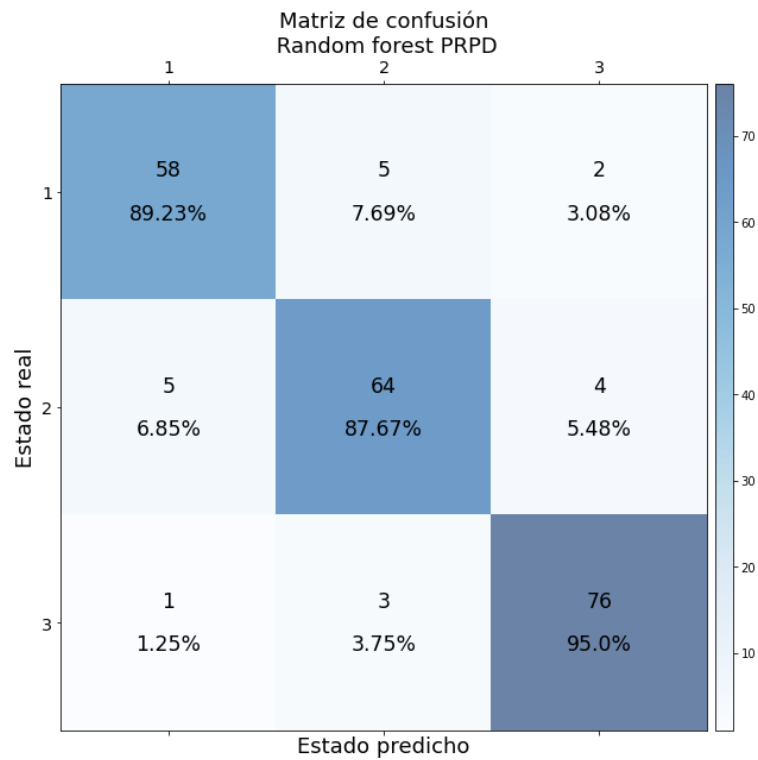


Figura 4.6: Matriz de confusión para random forest para características PRPD

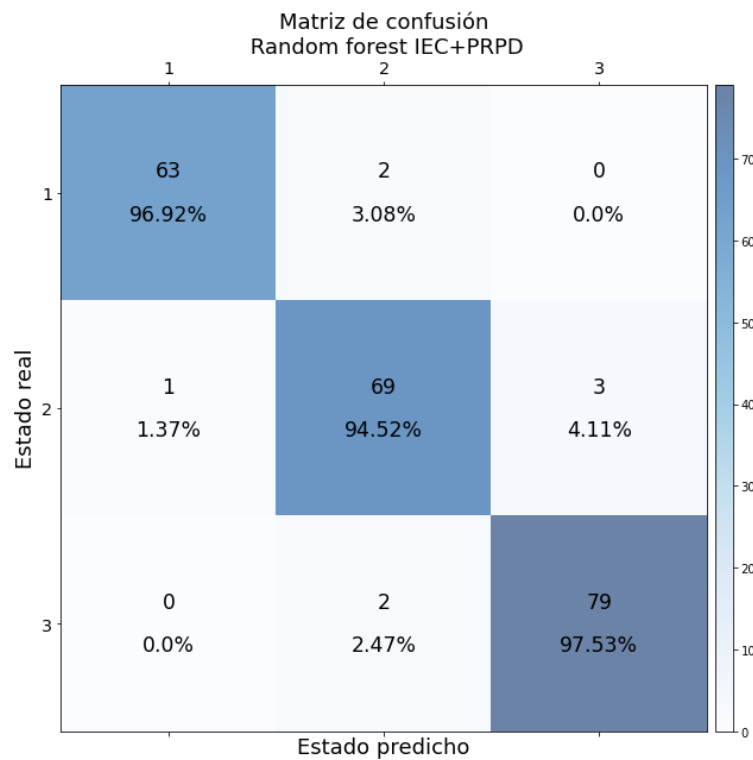


Figura 4.7: Matriz de confusión para random forest entrenado con características combinadas IEC y PRPD

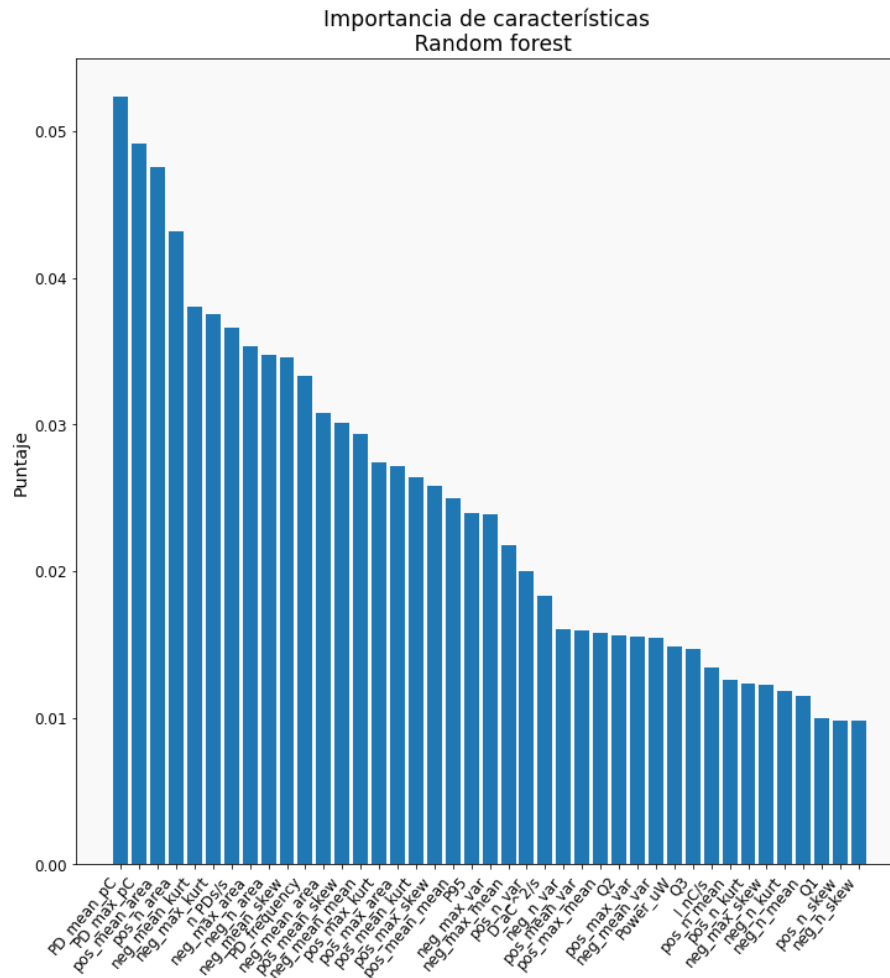


Figura 4.8: Importancia de características para random forest entrenado con características combinadas PRPD y IEC

4.3. Resultados redes neuronales

4.3.1. Predicción

La Tabla 4 muestra los resultados de predicción (exactitud en predecir la etapa) para los modelos de redes neuronales. Al igual que en el caso anterior, se desarrollaron tres modelos: uno que toma como dato de entrada solo características IEC, otro que tomo como dato de entrada solo características PRPD y el tercero que toma como dato de entrada ambos tipos de características. Se observa que para los modelos con una capa oculta utilizar ambas características, IEC y análisis PRPD, la exactitud no mejora respecto al caso donde se utiliza solo características PRPD, pero sí mejora respecto al caso IEC. Sin embargo, para el modelo con dos capas ocultas la exactitud es mayor para el caso donde solo se usan características PRDP respecto al caso donde se usan ambas características, IEC y PRPD, esto se puede explicar debido a que las redes neuronales tienden a crear sobreajuste a medida que aumentan las capas ocultas y el número de neuronas produciendo que no generalicen bien con datos no presentes en

el entrenamiento. Esto trae como consecuencia que al utilizar las características IEC y PRPD juntas la exactitud de la predicción disminuya. Para cada caso la exactitud es mejor para el modelo con dos capas ocultas respecto al modelo con una capa oculta utilizando las características IEC y PRPD tanto de forma independiente como conjunta.

Las matrices de confusión de la Figura 4.9 y Figura 4.10 muestran que hay una mejora para la menor exactitud por clase pasando de 86.36% a 88.64% indicando que el modelo con dos capas ocultas tiene una mejor capacidad para diferenciar entre los estados de propagación de árboles eléctricos. En general el modelo con dos capas ocultas es mejor tanto para la exactitud general como por clase. Además, se notó que, para la última etapa del árbol, siendo esta la más crítica por estar más cerca de la ruptura, el modelo no presentó error en su clasificación, lo que es prometedor puesto que evitaría una falla del equipamiento, es decir, el caso en que el deterioro este en etapa 3 y se prediga un estado diferente es improbable. Sin embargo, se observó que hay predicciones que indican un estado en etapa 3 cuando el árbol se encuentra en etapa 1 o 2, esto traería como resultado un diagnóstico erróneo que indica un estado crítico del aislante, sin este presentar un avance considerable del árbol. Los dos modelos de redes neuronales se comportaron en general de forma similar sin mostrar grandes diferencias en sus exactitudes.

La Figura 4.11 y Figura 4.12 muestra las épocas para los datos de entrenamiento y validación. En la sección 4.1, sobre los modelos propuestos, se mencionó que el modelo evaluado corresponde al de la época con el valor de costo de la función de entrenamiento (loss) menor. Esto se realizó así para evitar el sobreajuste hacia los datos de entrenamiento, esto se aprecia en los gráficos en un aumento constante del loss en los datos de validación, una disminución de la exactitud en la validación, y un aumento constante de la exactitud para la predicción sobre los datos de entrenamiento.

Con respecto a las características relevantes la Figura 4.13 muestra que para el modelo 2 las 5 características con mayor puntaje son PD_frequency, neg_n_kurt, pos_n_kurt, neg_mean_mean y neg_max_mean. De estas características la primera corresponde a características IEC y el resto a características PRPD. También se observa que de estas características las 3 primeras corresponden a la frecuencia de las descargas parciales, kurtosis de la distribución de cantidad de descargas parciales para el semiciclo positivo y la kurtosis para la distribución de cantidad de descargas parciales para el semiciclo negativo, indicando que es de importancia la fase de ocurrencia de los eventos de descargas parciales y la distribución estadística de esos tanto para el semiciclo positivo como para el negativo.

	Puntaje Exactitud		
	Tipo de característica		
Modelo red neuronal	IEC	PRPD	Todas
Modelo 1	0.8073	0.9174	0.9174
Modelo 2	0.8440	0.9450	0.9266

Tabla 4: Exactitud para modelos de red neuronal según tipo de características

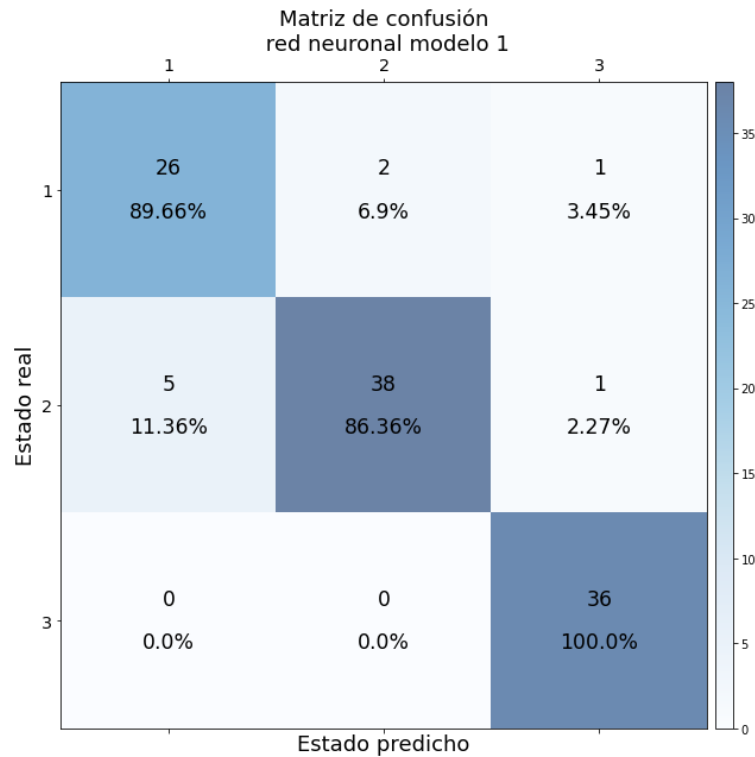


Figura 4.9: Matriz de confusión para red neuronal modelo 1 entrenada con todas las características

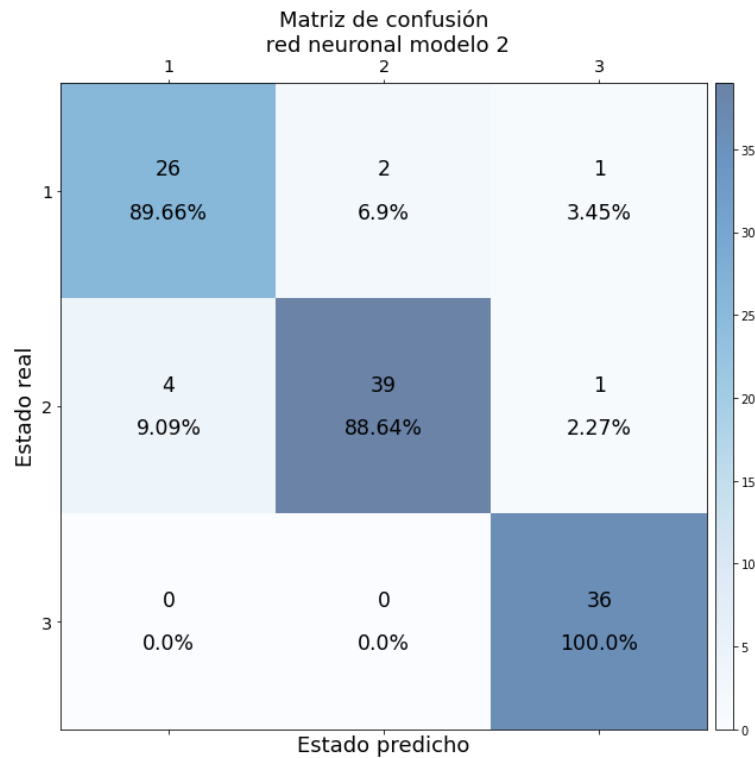


Figura 4.10: Matriz de confusión para red neuronal modelo 2 entrenada con todas las características

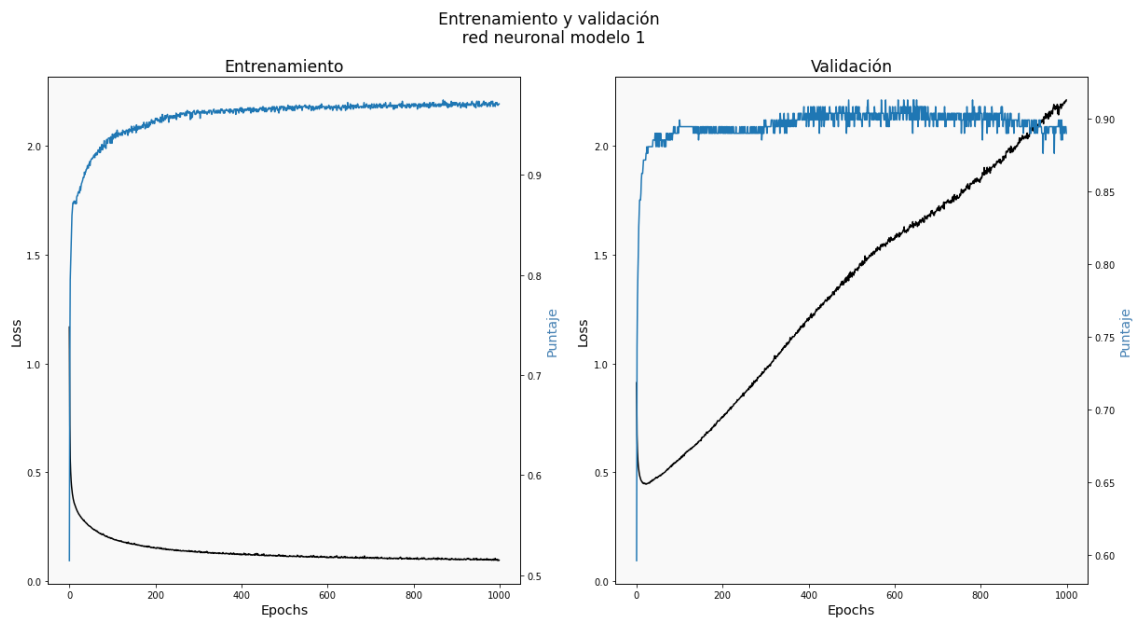


Figura 4.11: Épocas red neuronal modelo 1 para datos de entrenamiento y validación
IEC+PRPD

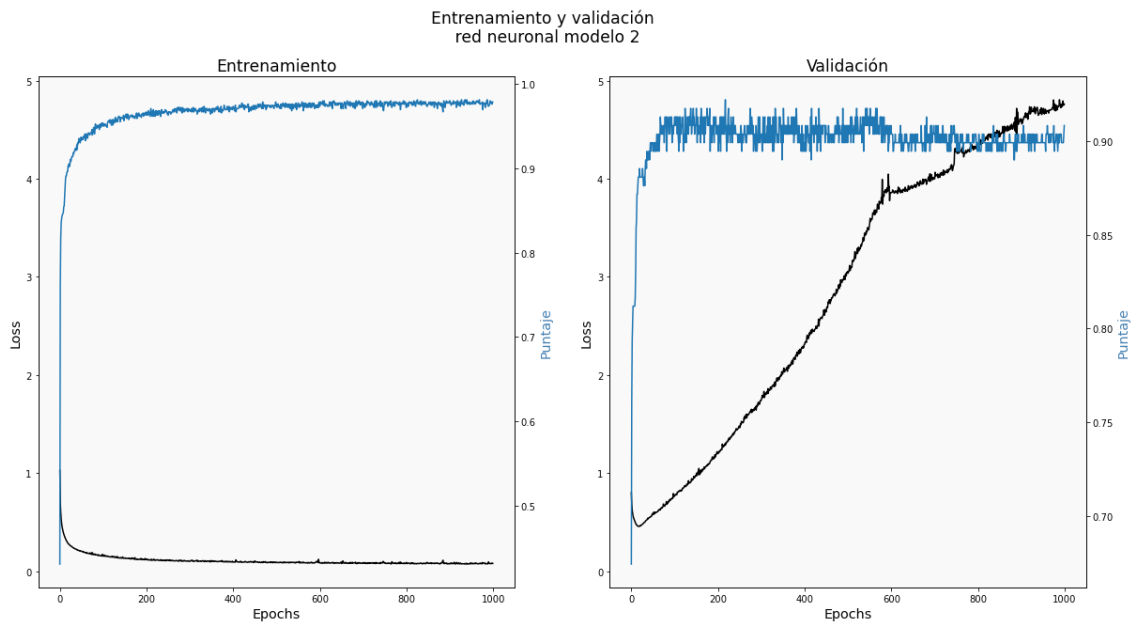


Figura 4.12: Épocas red neuronal modelo 2 para datos de entrenamiento y validación
IEC+PRPD

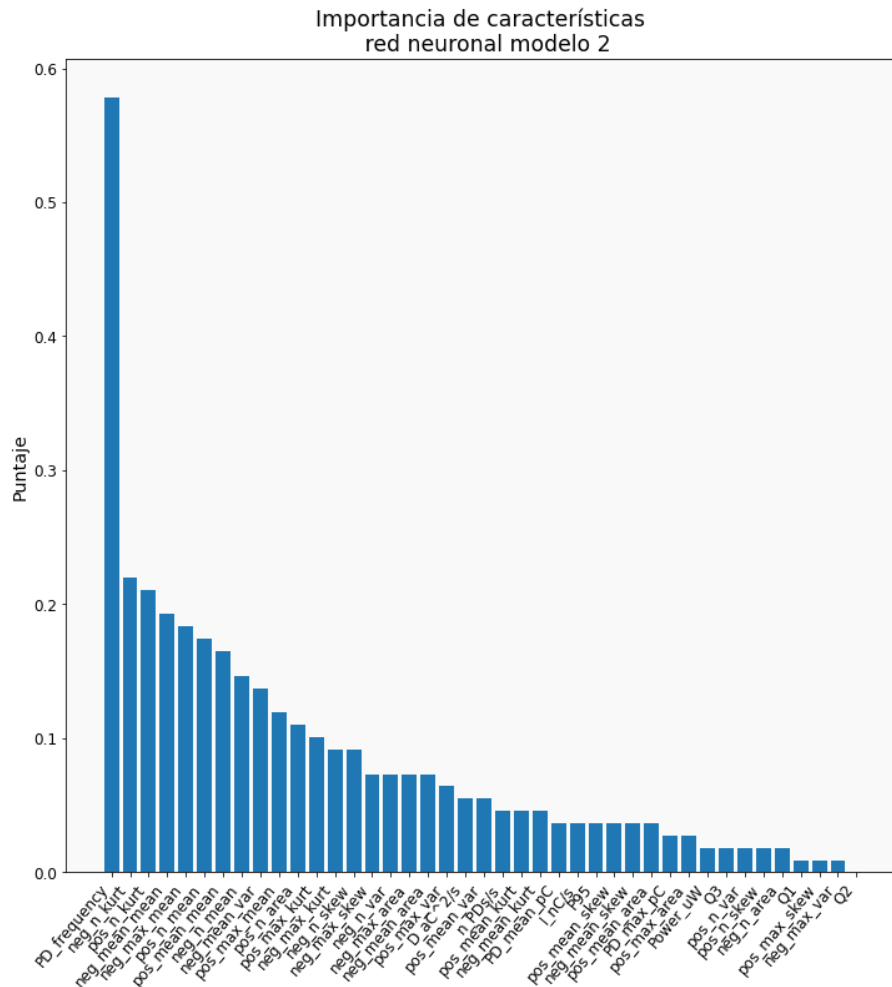


Figura 4.13: importancia de características red neuronal modelo 2

4.3.2. Ajuste de hiperparámetros red neuronal

Al ajustar la cantidad de neuronas y la cantidad de capas ocultas mediante búsqueda de grilla se obtuvo que la cantidad óptima corresponde a 1 capa ocultas con 10 neuronas cada una, llegando a una exactitud de 0.8945. Se debe hacer notar que a pesar de que esta no es la arquitectura con la mejor exactitud, si es la que tiene el menor costo de la función objetivo a minimizar. Esta optimización de hiperparámetros se realizó para observar si los modelos de redes neuronales propuestos coinciden con el óptimo. Sin embargo, los resultados obtenidos son no concluyentes debido a que los modelos presentan costos y exactitudes similares, pudiendo ser válidos para la predicción de las etapas de árboles eléctricos.

Neuronas	Capas	Exactitud	Valor de la función de costo
10	1	0.8853	0.5244
10	2	0.8853	0.6039
10	3	0.8761	0.7665
10	4	0.8945	0.6481
20	1	0.8853	0.5569
20	2	0.8945	0.6862
20	3	0.8761	0.6258
20	4	0.8624	0.9298
30	1	0.8853	0.5448
30	2	0.8945	1.0041
30	3	0.8991	1.0775
30	4	0.8899	0.8388
40	1	0.8991	0.5409
40	2	0.8991	0.929
40	3	0.8945	0.9736
40	4	0.9083	1.0775

Tabla 5: optimización de hiperparámetros red neuronal

4.4. Selección de características

La selección de características mediante eliminación recursiva para random forest mostrado en Figura 4.14 mostro una tendencia monótonamente creciente del puntaje respecto al número de características. Lo anterior podría indicar que todavía es posible mejorar aún más el modelo con la adición de más características relevantes. Una propuesta seria analizar la forma de onda de los pulsos de las descargas parciales mediante un análisis de espectro de frecuencias, como las descomposiciones wavelet presentadas en [29].

Para la eliminación recursiva de características para el modelo de red neuronal, mostrado en la Figura 4.16 se observa una tendencia monótonamente creciente del puntaje respecto al número de características con una alta variabilidad para una cantidad de características superior a 10. Esta variabilidad se puede asociar al sobreajuste y a la naturaleza estocástica del algoritmo de aprendizaje por retropropagación. El sobreajuste afectaría a esta variabilidad por el hecho de que el modelo estaría sobre ajustándose en para determinados entrenamientos.

Al analizar la importancia de características en ambos modelos se ve que hay características comunes dentro de las 14 mejores. Se ve la presencia de la cantidad de descargas en el intervalo de tiempo (PD_frequency) siendo esta la única característica IEC presente en ambos modelos,

donde la mayoría de las características son del tipo PRPD. Se observa también la presencia de las características de área generadas.

Lo anterior junto a lo expuesto en el análisis para redes neuronales estaría indicando que las características IEC no son tan relevantes como las de PRPD, por lo que, se podría prescindir de ellas a costa de perder un poco de exactitud en los modelos al igual como se mostró en la sección 4.3.

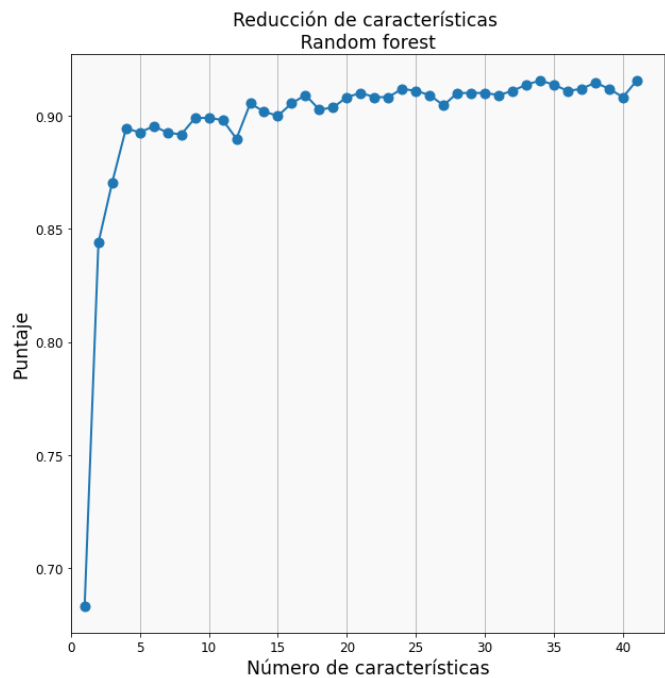


Figura 4.14: eliminación recursiva de características para random forest

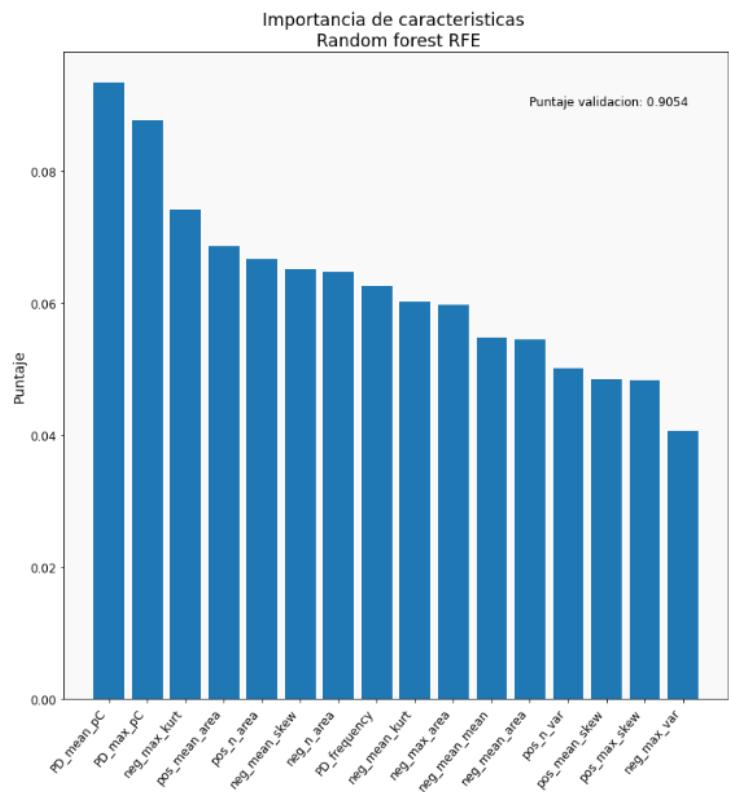


Figura 4.15: importancia para el caso de 15 características de la eliminación recursiva

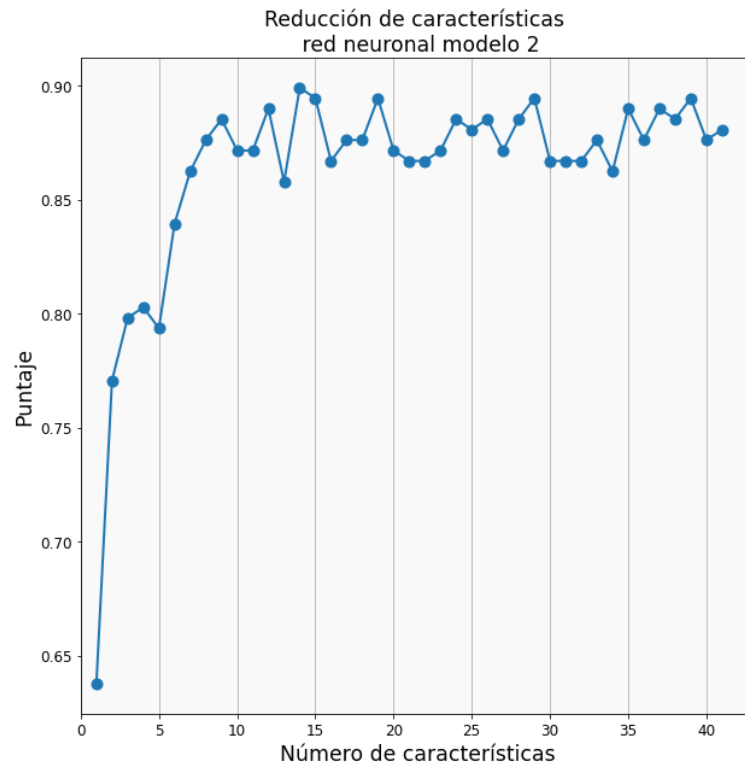


Figura 4.16: eliminación recursiva de características para red neuronal modelo 2

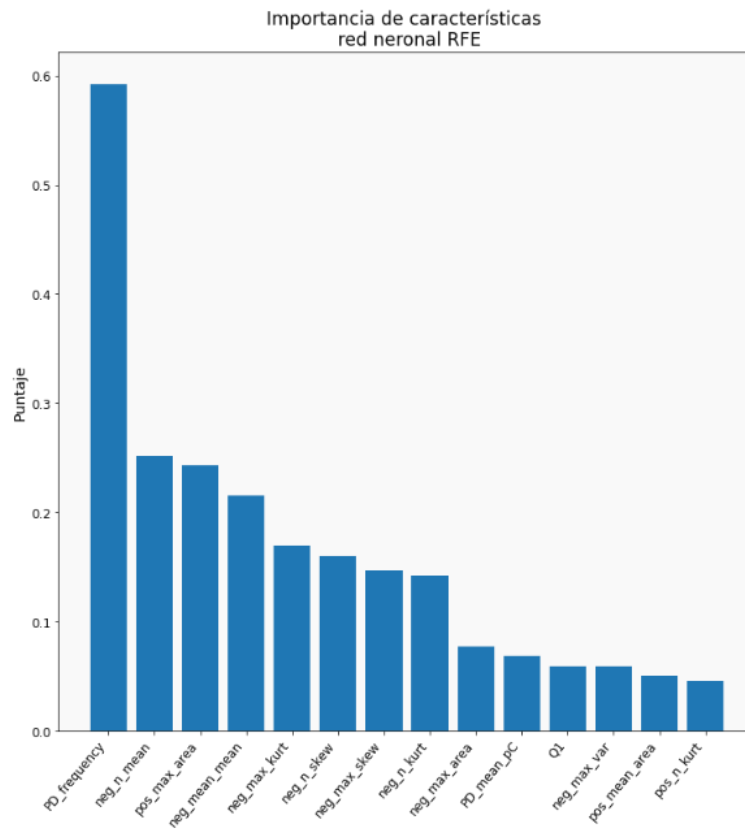


Figura 4.17: importancias para el caso de 14 características de eliminación recursiva.

Capítulo 5

Conclusiones

Se implementaron 3 modelos de inteligencia artificial correspondientes a random forest de 100 árboles de decisión, red neuronal de 1 capa oculta y una red neuronal de 2 capas ocultas clasificadores para evaluar la capacidad de predicción de la evolución de árboles eléctricos. Se entrenaron estos modelos de inteligencia artificial con datos caracterizados mediante parámetros IEC 60270 y diagramas resueltos en fase de descargas parciales. Los datos de entrenamiento se obtuvieron de probetas de resina epóxica ensayadas para 3 tipos de tensiones diferentes, siendo 15 kV AC, 15 kV AC + 15 kV DC y 15kV AC -15kV DC. Se definieron estados de deterioro para entrenar los modelos de inteligencia artificial, estos estados se definieron en base a las características de longitud y descargas parciales de los árboles eléctricos para cada probeta. Los modelos se entrenaron y evaluó su capacidad de predicción de forma independiente para las características IEC 60270 y los diagramas resueltos en fase para finalmente evaluarlos en forma conjunta. Además, se realizó una eliminación recursiva de características para el modelo de random forest y las redes neuronales de 2 capas ocultas.

Los modelos de random forest y redes neuronales de clasificación lograron exactitudes cercanas al 90%, similar a lo visto en la literatura, para la predicción del estado de propagación de árboles eléctricos. El modelo con la mejor capacidad de predicción corresponde a una exactitud del 92.66 por ciento, para el modelo 2 de redes neuronales con el set de datos combinados. Este resultado es prometedor en el sentido de ser un punto de partida para la evaluación preventiva de equipamiento eléctrico, más aún siendo los árboles un fenómeno de características estocásticas.

Se realizó un análisis sobre el tipo de características de cada modelo utilizando de forma independiente las características IEC y PRPD. Las características obtuvieron de forma independiente exactitudes entre 80% y 90% para los modelos de redes neuronales, y alrededor del 90 por ciento para random forest. Al combinar las características se obtuvieron mejoras entre 2 y 10 por ciento para los modelos. Además, se vieron mejoras en las exactitudes para cada clase en las matrices de confusión al combinar las características. Esto indica la adición de características aporta información complementaria a los modelos de redes neuronales y random forest, mejorando sus capacidades de predicción.

Para evaluar las características relevantes se realizó una eliminación recursiva, utilizando la importancia de Gini para random forest y la importancia de permutación para redes neuronales.

Estos métodos entregaron resultados similares respecto a la naturaleza de las características relevantes. Mostrando una dominancia de las características PRPD en todos los modelos. Los resultados sobre la cantidad de características mostraron tendencias crecientes de la exactitud respecto a la cantidad de características. Basado en esta tendencia se concluye que un modelo óptimo en el caso de la propagación de árboles eléctricos debe hacer uso de las características PRPD e IEC juntas. Se concluye también que las características PRPD son de suma importancia y no se debe prescindir de ellas al momento de entrenar los modelos, no es el caso de las características IEC las cuales se podrían omitir con una mínima reducción en la exactitud. Además, la tendencia creciente puede ser un indicador que todavía hay margen para obtener información sobre la propagación de árboles mediante la extracción de nuevas características, como por ejemplo complementar este trabajo con características de la forma de onda de los pulsos de las descargas parciales.

Se mostró que las técnicas de inteligencia artificial son efectivas al momento de predecir las propagaciones de árboles mediante un esquema de clasificación de estados. Esto coincide con las investigaciones para la predicción de la vida de equipamientos eléctricos a partir de estados de deterioro.

5.1. Trabajo a futuro

Se abre la posibilidad de realizar estudios sobre la predicción de estado de deterioro de aislantes mediante métricas sobre arboles eléctricos. También se puede complementar a futuro este trabajo con características de la forma de onda de los pulsos de descargas parciales, descomposición de wavelet, dimensión fractal de la forma de los árboles eléctricos, series no lineales de tiempo u otros, aportando más información a los modelos para evaluar su capacidad de predicción, apuntando a una mejora en la exactitud.

Referencias

- [1] J. Densley, «An Overview of Aging Mechanisms and Diagnostics for Extruded power cables,» de *2000 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.00CH37077)*, Singapore, 2000.
- [2] L. Dissado, «Understanding electrical trees in solids: from experiment to theory,» *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 9, nº 4, pp. 483-497, 2002.
- [3] C. Gao, Y. Yu, Z. Wang, W. Wang, L. Zheng y J. Du, «Study on the relationship between electrical tree development and partial discharge of XLPE cables,» *Journal of Nanomaterials*, 2019.
- [4] I. Iddrissu, *Study of electrical strength and lifetimes of polymeric insulation for DC applications*, 2016.
- [5] *High-voltage test techniques - partial discharge measurements*, 60270, IEC, 2000.
- [6] E. Kuffel, W. S. Zaengl y J. Kuffel, *High Voltage Engineering*, Great Britain: Newnes, 2000.
- [7] TECHIMP services srl, «Training course - Partial discharge diagnosis for rotating machines Part III: Diagnostic of rotating machines».
- [8] *IEEE Guide for Field Testing of shielded power cable systems rated 5kV and above with damped alternating current (DAC) voltage*, 400.4-2015, IEEE, 2016.
- [9] W. Wang, S. Chen, K. Yang, D. He y Y. Yu, «The relationship between electric tree aging degree and the equivalent time-frequency characteristic of PD pulses in high voltage cable,» de *2012 IEEE International Symposium on Electrical Insulation*, San Juan, PR, USA, 2012.
- [10] L. A. Dissado, S. J. Dodd, J. Champion, P. Williams y J. Alison, «Propagation of electrical tree structures in solid polymeric insulation,» *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 3, nº 3, pp. 259-279, 1997.
- [11] R. Schurch, P. Donoso, P. Aguirre, O. Cardenas, M. Zuniga y S. M. Rowland, «Electrical tree growth and partial discharges analyzed by fractal and correlation dimensions,» de *2017*

-
- IEEE Conference on Electrical Insulation and Dielectric Phenomenon (CEIDP)*, Fort Worth, TX, USA, 2017.
- [12] R. Shurch, O. Muñoz y J. Ardila-Rey, «Electrical tree growth under harmonic frequencies chracterized by partial discharge waveforms,» de *2020 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP)*, East Rutherford, NJ, USA, 2020.
- [13] R. Schurch, L. Orellan, P. Donoso-Daille, J. A. Ardila-Rey y J. Montana, «Pulse waveform, phase-resolved and pulse sequence analysis of partial discharges during electrical tree growth in epoxy resin,» de *Conference: International Symposium on High Voltage Engineering (ISH)*, Buenos Aires, Argentina, 2017.
- [14] H. Sui, L. Shen, W. Wang, Y. Yu, Z. Wang y C. Gao, «PD detecting during the growth of electrical trees in HV XLPE cable,» de *Asia-Pacific Power and Energy Engineering Conference*, Wuhan, China, 2011.
- [15] N. C. Sahoo, M. M. A. Salama y R. Bartnikas, «Trends in partial discharge pattern classification: a survey,» *IEEE*, vol. 12, nº 2, pp. 248-264, 2005.
- [16] L. Sun, S. Ji y J. Ye, *Multi-label dimensionality reduction*, Chapman and Hall, 2013.
- [17] V. Bolón Canedo, N. Sánchez Maroño y A. Alonso Betanzos, *Feature selection for high-dimensional data*, Springer Cham, 2015.
- [18] H. Liu y H. Motoda, *Feature extraction, construction and selection: A data mining perspective*, New York: Springer Science+Business Media, 1998.
- [19] S. Raschka, *Python machine learning*, Brimingham, United Kingdom: Packt publishing ltd., 2015.
- [20] C. M. Bishop, *Pattern recognition and machine learning*, Springer, 2006.
- [21] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, I. Goodfellow, A. Harp, G. Irving y M. Isard, «TensorFlow: Large-scale machine learning on heterogeneous systems,» 2015. [En línea]. Available: www.tensorflow.org. [Último acceso: Mayo 2023].
- [22] S. G. N., G. S. Sheshadri y N. C., «ANN Based Short Term Load Forecasting for Karnataka State Demand,» de *IEEE North Karnataka Subsection Flagship International Conference (NKCon)*, Vijaypur, India, 2022.

-
- [23] D. Mukherjee, S. Chakraborty, P. K. Guchhait y J. Bhunia, «Machine Learning based Solar Power Generation Forecasting with and without MPPT Controller,» de *IEEE 1st International Conference for Convergence in Engineering (ICCE)*, Kolkata, India, 2020.
- [24] M. Karthikeyan y R. Rengaraj, «Short-Term Wind Speed Forecasting Using Ensemble Learning,» de *7th International Conference on Electrical Energy Systems (ICEES)*, Chennai, India, 2021.
- [25] Z. Ahmed, J. V. Kluss y D. A. Wallace, «Partial discharge measurements and techniques for pattern recognition and life prediction of medium voltage XLPE cables,» de *2018 IEEE International Conference on High Voltage Engineering and Application (ICHVE)*, Atenas, Grecia, 2018.
- [26] S. K. S. P. Rakesh Sahoo, «Health index analysis of XLPE cable insulaion using machine learning technique,» de *IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, Prayagraj, India, 2020.
- [27] B. Liu, X.-h. Bao, J. Liu y Q.-l. Zhu, «Insulation life prediction of high voltage submersible motor based on BP neural network,» de *2011 International Conference on Consumer Electronics, Communications and Networks (CECNet)*, XiangNing, China, 2011.
- [28] S. H. Park, K. J. Lim y S. H. Kang, «Recognition of electrical tree progress in solid insulator,» de *2006 IEEE 8th International Conference on Properties & applications of Dielectric Materials*, Bali, Indonesia, 2006.
- [29] L. Li, Y. Wang, J. Gao, N. Guo y G. Han, «Pattern recognition of growth characteristics based ond UHF PD signals of electrical tree,» *IEEE Access*, vol. 10, pp. 49853-49861, 2022.
- [30] Ogura industrial jewels, «<http://www.ogura-indus.co.jp>,» [En línea]. Available: <http://www.ogura-indus.co.jp/en/product/section17.php>. [Último acceso: 18 Agosto 2023].
- [31] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot y E. Duchesnay, «Scikit-learn: Machine Learning in Python,» *Journal of Machine Learning Research*, vol. 12, pp. 2825--2830, 2011.
- [32] R. Shurch, P. Donoso, J. Ardila-Rey, J. Montana y A. Angulo, «Electrical tree growth under

very low frequency (VLF) voltage excitation,» de *2018 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP)*, Cancun, Mexico, 2018.

- [33] N. Roy, P. Tripathy, S. Chandra De, B. Swargiary, S. Kumar, S. Das y N. Pathak, «Day-ahead Solar Power Generation Forecasting using LSTM and Random Forest Methods for North Eastern Region of India,» de *22nd National Power Systems Conference (NPSC)*, New Delhi, India, 2022.

