

UNIVERSIDAD TÉCNICA DE AMBATO



FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E INDUSTRIAL MAESTRÍA EN FÍSICA APLICADA MENCIÓN FÍSICA COMPUTACIONAL

TEMA:

EVALUACIÓN CON RED NEURONAL DEL PROCESO DE DESGASTE
ABRASIVO DE PLACAS DE UN MATERIAL COMPUESTO DE LÁTEX CON
PARTÍCULAS DE CAUCHO RECICLADO.

Trabajo de titulación previo a la obtención del Grado Académico de Magister en Física
Aplicada mención Física Computacional

Modalidad De Titulación: Proyecto de Desarrollo

Autor: Ing. Pablo Rafael Muñoz Valverde

Director: Ing. Cristian Fabián Pérez Salinas, Mg.

Ambato – Ecuador

2022

APROBACIÓN DEL TRABAJO DE TITULACIÓN

A la unidad académica de titulación de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial. El Tribunal receptor del Trabajo de Titulación presidido por Ing. Elsa Pilar Urrutia Urrutia, Mg., e integrado por los señores: Ing. Víctor Hugo Hidalgo Diaz, Dr. y Fís. Christian Leonardo Vásquez Vega, Dr., designados por la Unidad de Titulación de la Universidad Técnica de Ambato, para receptor el Trabajo de Titulación con el tema: “Evaluación con red neuronal del proceso de desgaste abrasivo de placas de un material compuesto de látex con partículas de caucho reciclado”, elaborado y presentado por el señor Ingeniero, Pablo Rafael Muñoz Valverde, para optar por el Grado Académico de Magister en Física Aplicada mención Física Computacional; una vez escuchada la defensa oral del Trabajo de Titulación el Tribunal aprueba y remite el trabajo para uso y custodia en las bibliotecas de la Universidad Técnica de Ambato.

Ing. Elsa Pilar Urrutia Urrutia, Mg.
Presidente y Miembro del Tribunal de Defensa

Ing. Víctor Hugo Hidalgo Diaz, Dr.
Miembro del Tribunal de Defensa

Fís. Christian Leonardo Vásquez Vega, Dr.
Miembro del Tribunal de Defensa

AUTORÍA DEL TRABAJO DE TITULACIÓN

La responsabilidad de las opiniones, comentarios y críticas emitidas en el trabajo de titulación presentado con el tema: “Evaluación con red neuronal del proceso de desgaste abrasivo de placas de un material compuesto de látex con partículas de caucho reciclado”, le corresponde exclusivamente a: Ingeniero, Pablo Rafael Muñoz Valverde, autor bajo la dirección del Ingeniero Cristian Fabián Pérez Salinas, Mg., director del trabajo de titulación; y el patrimonio intelectual pertenece a la Universidad Técnica de Ambato.

Ing. Pablo Rafael Muñoz Valverde

AUTOR

Ing. Cristian Fabián Pérez Salinas, Mg.

DIRECTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que el Trabajo de titulación, sirva como un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la Institución.

Cedo los Derechos de mi trabajo, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad Técnica de Ambato.

Ing. Pablo Rafael Muñoz Valverde

1804374088

ÍNDICE GENERAL

CONTENIDO	
PORTADA.....	i
APROBACIÓN DEL TRABAJO DE TITULACIÓN	ii
AUTORÍA DEL TRABAJO DE INVESTIGACIÓN.....	iii
DERECHOS DE AUTOR	iv
ÍNDICE DE TABLAS	viii
ÍNDICE DE FIGURAS.....	ix
AGRADECIMIENTO	xiii
DEDICATORIA	xiv
RESUMEN EJECUTIVO.....	xv
EXECUTIVE SUMMARY.....	xvi
CAPÍTULO I.....	1
EL PROBLEMA DE INVESTIGACIÓN.....	1
1.1. Introducción	1
1.2. Justificación	2
1.3. Objetivos.....	3
1.3.1. Objetivo General:	3
1.3.2. Objetivos Específicos:.....	3
CAPITULO II	4
ANTECEDENTES INVESTIGATIVOS	4
2.1. Antecedentes Investigativos	4
2.2. Fundamentación teórica.....	7
2.2.1. Tribología.....	7
2.2.2. Machine Learning (ML).....	10
2.2.3. Red neuronal artificial (RNA).....	10

2.2.4. Modelo de Perceptrón Multicapa.....	14
2.2.5. Función de activación	16
2.2.6. Backpropagation	20
2.2.7. Algoritmos de los Optimizadores.....	24
2.2.8. Función de pérdida.....	29
CAPITULO III	31
MARCO METODOLÓGICO	31
3.1. Ubicación.....	31
3.2. Equipos y materiales.....	31
3.3. Tipo de investigación.....	31
3.3.1. Investigación bibliográfica.....	31
3.3.2. Investigación aplicada.....	32
3.3.3. Investigación cuantitativa.....	32
3.4. Prueba de hipótesis.	32
3.5. Población o muestra.....	33
3.6. Recolección de información.	33
3.7. Procesamiento de la información y análisis estadístico.....	34
3.8. Variables de la investigación	35
CAPITULO IV	37
RESULTADOS Y DISCUSIÓN	37
4.1. Densidad y volumen del látex sintético	39
4.2. Densidad y volumen del gránulo de caucho reciclado	41
4.3. Ensayos de desgaste abrasivo	47
4.1. Recolección de los resultados obtenidos del ensayo de desgaste abrasivo....	49
4.4. Desarrollo de la Red neuronal artificial	55
4.4.1. Función de activación y optimizador	59
4.4.2. Tasa de aprendizaje y Batch size	62
4.4.3. Épocas y sobreajuste	65
4.5. Análisis de resultados	69
4.6. Discusión	75

CAPITULO V78

CONCLUSIONES Y RECOMENDACIONES78

5.1. Conclusiones78

5.2. Recomendaciones79

5.3. Bibliografía80

ÍNDICE DE TABLAS

Tabla 1: Diferencias entre la Inteligencia Artificial y Machine Learning	10
Tabla 2: Conjunto de datos	33
Tabla 3: Recolección de información	34
Tabla 4: Variables de estudio.....	36
Tabla 5: Dimensiones del molde.....	38
Tabla 6: Mediciones de la masa de látex sintético.....	40
Tabla 7: Mediciones del volumen del látex sintético.....	40
Tabla 8: Mediciones del volumen del granulado de caucho.....	41
Tabla 9: Densidad de matriz y refuerzo	42
Tabla 10: Volumen y masa de las composiciones volumétricas.....	44
Tabla 11: Denominación de las composiciones volumétricas	44
Tabla 12: Grupo de identificación de probetas para recolección de datos.....	45
Tabla 13: Procedimiento para las pruebas de desgaste abrasivo.....	47
Tabla 14: Recolección de datos con la configuración de 60 % Matriz – 40% PGCR	49
Tabla 15: Recolección de datos con la configuración de 70 % Matriz – 30% PGCR	51
Tabla 16: Recolección de datos con la configuración de 70 % Matriz – 30% PGCR	53
Tabla 17: Conjunto de datos de entrada.....	56
Tabla 18: Validación estadística conforme a diferentes configuraciones neuronales.....	68
Tabla 19: Resultados estadísticos de MAE, MSE y R^2 de modelo neuronal.....	70
Tabla 20: Hiperparámetros del modelo neuronal.....	71

ÍNDICE DE FIGURAS

Figura 1: Fricción producida por dos superficies rugosas colocadas en contacto. W corresponde al peso de la superficie superior. Fuente: [19].....	7
Figura 2: Mecanismos de desgaste. En el panel a) se muestra el desgaste por adhesión entre dos superficies deslizantes. En el panel b) se presenta el desgaste por abrasión entre una superficie dura y una más blanda. En el panel c) se muestra el desgaste por fatiga entre dos materiales por cargas cíclicas. En el panel d) se observa el desgaste por corrosión causado por la formación de una película en la superficie del material. Fuente: [42]	8
Figura 3: Formas de desgaste abrasivo. En el caso (a) se muestra el desgaste por rugosidad o picos de una superficie directamente sobre otra superficie blanda. En el caso (b) se presenta el desgaste entre dos superficies en contacto con partículas abrasivas libres. Fuente: [19]	9
Figura 4: Estructura de una red neuronal biológica conformado por dendritas que transmiten estímulos al núcleo, donde se procesa la información y se envía una respuesta a través del axón. Fuente: [6]	11
Figura 5: Modelo neuronal artificial conformado por la capa de entrada, capas ocultas y una capa de salida. Cada nodo se conecta con otro y tiene asociado un peso y un umbral, finalmente la capa de salida proporciona la predicción final.....	12
Figura 6: Modelo neuronal de aprendizaje supervisado basado en el reajuste de pesos de conexión de cada neurona controlado por un agente externo, y determina la respuesta de salida de la red.....	14
Figura 7: Modelo neuronal Perceptrón Multi-Capa compuesto por la capa de entrada, capa oculta y una capa de salida con $k = 1$	15
Figura 8: Gráfica de la función escalón con un umbral igual a 0.....	17
Figura 9: Gráfica de la función lineal como función de activación.	17
Figura 10: Gráfica de la función sigmoidea como función de activación.....	18
Figura 11: Gráfica de la función de activación de la tangente hiperbólica representada por la línea naranja y la función Sigmoidea de color azul.	19
Figura 12: Gráfica de la función de activación de una ReLU con valores positivos a	

partir de cero, se caracteriza por no presentar valores negativos.....	19
Figura 13: Gráfica de la función de activación Leaky ReLU con una pendiente para los valores negativos, evitando que las neuronas se desactiven.	20
Figura 14: Diagrama del algoritmo de backpropagation. La primera fase consiste en el Feedforward el cual envía las señales hacia delante, los resultados se controlan a través del error de tolerancia. Si los resultados no presentan un valor satisfactorio se aplica el backpropagation, que consiste en reajustar los pesos y bias de las neuronas de la red. ..	21
Figura 15: Método de descenso de gradiente utilizado para minimizar una función de pérdida y converger hacia el valor mínimo de una función mediante un proceso iterativo. Fuente: [46]	25
Figura 16: Método de momentum frente a gradiente descendiente. La línea roja representa el algoritmo de momentum diseñado para acelerar el proceso de optimización mediante impulsos, y las flechas negras al gradiente descendiente basado en la información del gradiente anterior hasta encontrar el mínimo local. Fuente: [46].....	27
Figura 17: Diagrama del proceso metodológico de la investigación.	35
Figura 18: Diseño CAD del molde de las probetas de acuerdo con la norma ISO 9352, conformado por 4 cavidades homogéneas.	38
Figura 19: Molde mecanizado en la maquina CNC de la empresa "ColorPlast"	39
Figura 20: Proceso de estratificación manual de las probetas para encontrar una composición homogénea.	44
Figura 21: Proceso de desgaste abrasivo sobre una placa de ensayo.	48
Figura 22: El coeficiente de correlación que mejor valor presenta fue entre la relación de IDT y el peso aplicado con una correlación de 0,74, mientras que la correlación más baja corresponde al IDT y los ciclos de desgaste con un valor de 0,15.....	57
Figura 23: Distribución muestral de la variable de salida IDT caracterizada por una forma curva asimétrica positiva respecto a un valor central que coincide con la media y mediana de la distribución, la cual se extiende en la dirección positiva.....	57
Figura 24: En el apartado (a) se muestra el error de entrenamiento y validación con los datos sin normalizar, el cual presenta muchas oscilaciones. En el apartado (b) se utilizó los datos normalizados, mejorando su estabilidad y desempeño.	58
Figura 25: Error de entrenamiento y validación con el algoritmo de optimizado	

RMSprop. (a) ReLu desempeño buenos resultados con bajas oscilaciones. (b) Sigmoid presentó muchas oscilaciones durante todo el entrenamiento sin llegar a converger. (c) Tanh demostró oscilaciones muy bruscas durante todo el proceso. (d) LeakyReLu obtuvo buenos resultados, pero con oscilaciones pequeñas.....60

Figura 26: Error de entrenamiento y validación con el algoritmo de optimizado Adam. (a) ReLu demostró un buen desempeño y bajas oscilaciones. (b) Sigmoid empieza a converger a partir de la época 100. (c) Tanh presentó bajas oscilaciones, pero muy inestables. (d) LeakyReLu obtuvo oscilaciones más grandes con respecto a ReLU.60

Figura 27: Error de entrenamiento y validación con SGD. (a) ReLu presenta buenos resultados con pocas oscilaciones al inicio del entrenamiento. (b) Sigmoid converge hasta la época 100 con valores elevados. (c) tanh muestra grandes oscilaciones durante el entrenamiento. (d) LeakyReLu obtuvo resultados similares a ReLu.61

Figura 28: Error de entrenamiento y validación con $BS = 1$. (a) Presentó oscilaciones durante todo el entrenamiento. (b) y (c) mostraron valores muy elevados para el error de entrenamiento, dejándolos fuera del análisis.62

Figura 29: Error de entrenamiento y validación con $BS = 13$. (a) Presentó muy pocas oscilaciones y un buen desempeño. (b) y (c) mostraron valores elevados para el error de entrenamiento.63

Figura 30: Error de entrenamiento y validación con $BS = 32$. (a) Presentó pocas oscilaciones y un buen desempeño. (b) y (c) mostraron valores altos para el error de entrenamiento.64

Figura 31: Modelo con overfitting, donde el error de entrenamiento disminuye mientras el error de validación aumenta.65

Figura 32: Análisis de overfitting del modelo neuronal. En el apartado (a) se realizó el entrenamiento sin utilizar EarlyStopping, demostrando que el número de épocas llegó a su máximo de 1000, mientras que el apartado (b) utilizando EarlyStopping el entrenamiento se detuvo en 520 épocas aproximadamente.66

Figura 33: Análisis de underfitting del modelo neuronal. (a) Al utilizar el 20% de los datos el error de validación supera el error de entrenamiento. (b) Al entrenar el modelo con el 50% disminuye el error, pero no presenta una buena predicción. (b) Presenta el mejor error de entrenamiento con el 80% de los datos.67

Figura 34: Validación cruzada de K-fold para evaluar el rendimiento del modelo mediante el entrenamiento de k subconjuntos de datos y evaluados en el subconjunto complementario. Fuente: [44]	69
Figura 35: Regresión entre los datos presentados por la red neuronal en comparación con los valores reales obtenidos de manera experimental.	70
Figura 36: Error de entrenamiento y validación de la red. (a) Representa el MAE obtenido por la red, donde se obtuvo un valor de 0,749 mientras que el error de validación fue de 1,198. (b) Representa el MSE obteniendo por la red, presentando un valor de 1,607 y el error de validación se obtuvo 3,69.	71
Figura 37: La grafica de distribución de residuos tiende a una curva gaussiana, es decir, la mayoría de los residuos tienden a acercarse al valor de cero.....	72
Figura 38: Datos predichos por el modelo neuronal toman la forma de los datos obtenidos de manera experimental.....	72
Figura 39: Porcentaje del error de predicción en comparación con los valores reales. ...	73
Figura 40: La correlación de Pearson entre variables presenta una relación positiva entre el IDT y el peso, mientras que la peor la relación corresponde al porcentaje volumétrico y el peso aplicado.	74
Figura 41: La dispersión de residuos distribuidos de forma homogénea. El patrón de residuos se distribuye al azar, no presenta tendencias ni tampoco dispersiones.....	74
Figura 42: Gráfico de cuantiles teóricos del modelo. Los residuos presentan una distribución normal a lo largo de la línea de tendencia.....	75

AGRADECIMIENTO

Agradezco a mi madre Janet Valverde y a mi hermana Grace Muñoz por el apoyo constante para generar mi desarrollo personal y profesional, siendo el pilar y el motor de todas mis acciones.

A mi madre por que participó de este duro y difícil proceso y siempre me dio ánimo para seguir adelante.

A mi director de tesis y revisores por la paciencia y la entrega para generar este trabajo con la mayor calidad posible.

Pablo

DEDICATORIA

Dedicado a mi madre y mi hermana,
por ser la razón de mi trabajo constante;
por ser parte de este proceso de
desarrollo personal y profesional.

A mi madre, por inculcar en mí los
valores y la responsabilidad para ser
cada día una mejor persona. Pero sobre
todo por la confianza depositada en mi
para culminar este proceso.

Pablo

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL

MAESTRÍA EN FÍSICA APLICADA MENCIÓN FÍSICA COMPUTACIONAL

TEMA: “Evaluación con red neuronal del proceso de desgaste abrasivo de placas de un material compuesto de látex con partículas de caucho reciclado”

AUTOR: Ing. Pablo Rafael Muñoz Valverde

DIRECTOR: Ing. Cristian Fabián Pérez Salinas, Mg.

LÍNEA DE INVESTIGACIÓN:

- Materiales y procesos de fabricación para la industria.

FECHA: 30 de noviembre del 2021

RESUMEN EJECUTIVO

En la presente tesis se investigó un enfoque de Machine Learning en el campo de la fabricación de nuevos materiales para la industria. En particular, se utilizaron redes neuronales artificiales para predecir el índice de desgaste Taber (IDT) de placas de látex y partículas de caucho reciclado. En los últimos años ha aumentado sustancialmente la aplicación de la Inteligencia Artificial y en particular del Machine Learning a las disciplinas científicas. El propósito fue evaluar el funcionamiento del aprendizaje automático, en particular las redes neuronales, y cómo debe aplicarse para realizar una predicción. La fase preliminar del trabajo fue crear el conjunto de datos obtenidos de manera experimental necesarios para la fase secundaria, el cual abarca el análisis y modelado de redes neuronales. La generación del conjunto de datos involucró la fabricación del material y pruebas de desgaste en base a la norma ISO 9352. En el contexto de la red neuronal, se utilizó el software Google TensorFlow a través de la interfaz de Python3. El modelo desarrollado permite predecir el IDT de la placa tomando como variables independientes; el porcentaje volumétrico de material, la velocidad de giro, carga aplicada y el número de ciclos. El rendimiento de la red se evaluará mediante pruebas estadísticas como el error cuadrático medio (MSE), el error absoluto medio (MAE) y el coeficiente de determinación(R^2).

Palabras clave: Machine Learning, Inteligencia Artificial, Red neuronal artificial, índice de desgaste taber, predicción, TensorFlow, Python3, MSE, MAE, R^2 .

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL

MAESTRÍA EN FÍSICA APLICADA MENCIÓN FÍSICA COMPUTACIONAL

THEME: " Neural network evaluation of the abrasive wear process of plates made of a latex composite material with recycled rubber particles "

AUTHOR: Ing. Pablo Rafael Muñoz Valverde

DIRECTED BY: Ing. Cristian Fabián Pérez Salinas, Mg.

LINE OF RESEARCH:

- Materials and manufacturing processes for industry.

DATE: November 30th, 2021

EXECUTIVE SUMMARY

In this thesis, a Machine Learning approach was investigated in the field of manufacturing new materials for industry. In particular, artificial neural networks were used to predict the Taber wear index (TDI) of latex plates and recycled rubber particles. In recent years, the application of Artificial Intelligence and in particular Machine Learning to scientific disciplines has increased substantially. The purpose was to evaluate how machine learning works, in particular neural networks, and how it should be applied to make a prediction. The preliminary phase of the work was to create the experimentally obtained data set necessary for the secondary phase, which includes the analysis and modeling of neural networks. The generation of the data set involved the manufacture of the material and wear tests based on the ISO 9352 standard. In the context of the neural network, the Google TensorFlow software was used through the Python3 interface. The model developed allows to predict the IDT of the plate taking as independent variables; the volumetric percentage of material, the rotational speed, applied load and the number of cycles. The performance of the network will be evaluated through statistical tests such as the mean square error (MSE), the mean absolute error (MAE) and the coefficient of determination (R²).

Keywords: Machine Learning, Artificial Intelligence, Artificial neural network, taber attrition index, prediction, TensorFlow, Python3, MSE, MAE, R².

CAPÍTULO I

EL PROBLEMA DE INVESTIGACIÓN

1.1. Introducción

Según un estudio de la Unión Internacional para la Conservación de la Naturaleza (UICN), las micropartículas de plástico que se desprenden de productos industriales como ropa sintética, neumáticos entre otros. Estos materiales representan hasta el 30% de la basura que contaminan los océanos y ha aumentado significativamente en los últimos años [1]. En Ecuador se desechan 2.4 millones de neumáticos en el año, y solo el 1% de ese valor es utilizado para el reencauche, el resto de los neumáticos se lo quema y provoca la emisión de sustancias químicas a la atmósfera que afectan directamente a la calidad del aire [2]. En la actualidad, existen varios métodos para reutilizar el desecho de llantas u otros elementos fabricados con materiales elastómeros, ya sea mediante la transformación total o parcial de los mismos [3]. En Ecuador, existe poco conocimiento de productos fabricados con materia prima a base de partículas de caucho reciclado. Esto provoca un problema ambiental si no se los trata de manera adecuada, dado que estos materiales no pueden regresar al medio ambiente de manera ecológica a través de la degradación biológica, la hidrolización o la descomposición. Proceso que tarda alrededor de 500 años en degradarse y, si son incineradas, desprenden grandes cantidades de metano, un gas de efecto invernadero [4].

El aprendizaje automático (Machine Learning) y la Inteligencia Artificial (IA) son hoy en día de gran interés en varios campos científicos que contribuyen al medio ambiente [5]. La inteligencia artificial se puede aplicar a varios sectores diferentes de la industria, por ejemplo, la predicción de dureza, inflamabilidad, mejorar la eficiencia operativa, el diseño de productos, el servicio al cliente [6]. El aprendizaje automático se puede utilizar para manejar una gran cantidad de datos y extraer conocimiento de posibles nuevos materiales. El aprendizaje automático proporciona un marco que se puede adaptar para abordar muchos desafíos en la industria, como el procesamiento de datos experimentales, la optimización de material, el modelado y el control de la calidad [7].

En el presente proyecto se investigó técnicas de aprendizaje automático y su aplicabilidad a la predicción del desgaste abrasivo en placas de látex y partículas de caucho reciclado, con diferentes combinaciones volumétricas y diferentes condiciones de desgaste. El proyecto está organizado de la siguiente manera: En el capítulo 2 se proporciona los antecedentes investigativos en torno al desgaste por abrasión y fundamentación teórica de las redes neuronales. En el capítulo 3 se detalla la metodología de estudio. El capítulo 4 presenta las pruebas realizadas para la generación del conjunto de datos en base a la norma ISO 9352 y se analizarán los resultados obtenidos luego del entrenamiento de la red, mediante pruebas estadísticas. Por último, las conclusiones y recomendaciones se establecen en el capítulo 5.

1.2. Justificación

En la actualidad, Ecuador genera cada año 250 000 toneladas de cauchos, de los que aproximadamente la mitad se depositan en vertederos controlados sin tratar [8]. La práctica general para la eliminación de estos residuos ha sido la quema directa, una medida que provoca graves problemas medioambientales y emisiones de gases que contienen partículas nocivas para el entorno. Estas medidas se han convertido en una preocupación global por su impacto negativo en el medio ambiente y en la salud humana [9].

Los análisis tribológicos actuales, permiten conocer el comportamiento de materiales en un proceso industrial, aunque algunas de estas propiedades son por regla general constantes, se ha demostrado que estas características se ven modificadas de forma espontánea por diversos factores, tales como: los mecanismos de desgaste, temperatura, regímenes de lubricación, entre otros. Este comportamiento se considera como no lineal, contrario a lo sugieren los postulados clásicos del desgaste [10]. Un modelo clásico de desgaste planteado por Archard permite realizar una proyección para predecir el desgaste mecánico de forma lineal. Sin embargo, las pruebas experimentales han demostrado no linealidad en el proceso de desgaste mecánico [11]. La generación de materiales compuestos a base de partículas de caucho es una posible solución a la contaminación ambiental provocado por estos elastómeros. Estos compuestos resultan de combinar dos

o más materiales, denominados matriz y refuerzo [12]. En esta investigación se ha seleccionado como matriz al látex y como refuerzo el gránulo de caucho reciclado, obtenido luego de un proceso de trituración de llantas de vehículos. Reutilizando este material se colaborará con el reciclaje y disminuirá los niveles de contaminación.

El potencial de las redes neuronales artificiales (RNA) radica en su naturaleza flexible y su capacidad para capturar relaciones no lineales complejas entre los patrones de entrada y salida. El aprendizaje automático es aplicado a un conjunto de datos representativos que debe contener información suficiente del problema a tratar [13]. Las redes neuronales presentan una alta tolerancia a fallos del sistema y proporcionan un alto grado de paralelismo en el proceso de datos. Por esta razón, es posible implementar modelos neuronales de bajo costo [14]. En la presente investigación se plantea un modelo de red neuronal capaz de predecir el índice de desgaste abrasivo de las placas mencionadas anteriormente, y será evaluada mediante análisis estadístico, a fin de determinar la precisión del modelo, así como dar un rango de confianza de la estimación, determinando de esta forma su posible inmersión en la industria.

1.3. Objetivos

1.3.1. Objetivo General:

1. Evaluar el proceso de desgaste abrasivo de placas obtenidas a partir de un material compuesto de látex con partículas de caucho reciclado mediante una red neuronal artificial.

1.3.2. Objetivos Específicos:

1. Realizar pruebas de desgaste abrasivo en las placas del material planteado para cuantificar el desgaste experimental.
2. Elegir y desarrollar la arquitectura óptima de la red neuronal artificial para la predicción del desgaste abrasivo.
3. Evaluar el rendimiento del modelo de red mediante pruebas estadísticas.

CAPITULO II

ANTECEDENTES INVESTIGATIVOS

2.1. Antecedentes Investigativos

En los últimos años, la tribología se ha convertido en uno de los campos más importante de la ingeniería. En esta ciencia se considera los efectos del movimiento, la interacción entre el material, el lubricante y la herramienta o material de revestimiento durante el proceso de fricción [15]. El desafío en este campo es representar los fenómenos tribológicos complejos a través de un software fácil de utilizar [16]. Este problema es debido a que se utilizan muy pocas variables para caracterizar el sistema de estudio, entre ellas están: las propiedades mecánicas y geométricas del material, la lubricación, propiedades elásticas de sólidos y problemas de tensión de contacto [17], [18]. El desgaste que sufren los materiales es un daño superficial inducido mecánicamente provocando como resultado la eliminación progresiva del material debido al movimiento relativo entre la superficie y el material en contacto [19]. El desgaste abrasivo es la pérdida de material debido a partículas o protuberancias duras que se fuerzan y se mueven a lo largo de una superficie sólida. Esto ocurre cuando una superficie dura y rugosa se desliza sobre una superficie más blanda [20]. En la mayoría de las aplicaciones tecnológicas, la aparición de desgaste es altamente indeseable, provocando el deterioro o incluso al fallo de los componentes [21].

Los metales siempre han sido la primera opción para el diseño de ingeniería desde que se inventó el acero. Los materiales compuestos surgen de la necesidad de fabricar elementos con altas propiedades mecánicas específicas, alta relación rigidez/peso, alta relación resistencia/peso, adaptabilidad y tolerancia al daño [12]. Están formados por la combinación de dos o más constituyentes distintos. Las principales ventajas son: la reducción de peso, la mejora de la resistencia al desgaste y la corrosión, el aislamiento térmico acústico, la gran resistencia específica, la vida útil mejorada y la baja expansión térmica [22]. El proceso de desgaste generalmente se verifica experimentalmente, ya que los parámetros del proceso como los ciclos de deslizamiento, la carga aplicada, la

velocidad, las cualidades del material y la textura de la superficie influyen en la calidad de los ensayos. Los fenómenos de superficie son complejos y su modelado matemático aún es limitado. En los últimos años, el uso de la tecnología basada en inteligencia artificial (IA) se ha vuelto popular en los estudios tribológicos. El-Aziz [23] estudió la tasa de desgaste abrasivo de hierros fundidos blancos con alto contenido de Cr (HCCI) mediante técnicas de machine learning. Se obtuvo un error cuadrático medio de 4.76. En 2016, Radhakrishnan [13] desarrolló un modelo neuronal para predecir las características de desgaste del compuesto Al 2219-SiC. Se presentó una buena precisión de predicción de aproximadamente el 85%. Kopal [24] investigó la respuesta de tracción uniaxial de mezclas de caucho natural vulcanizado con diferentes contenidos de carbono. LiuJie y Cardoso [25] utilizaron modelos neuronales para estudiar los efectos del factor pv (producto de la presión y la velocidad) y la temperatura de contacto en el comportamiento tribológico de deslizamiento en seco del compuesto de polieteretercetona reforzado con fibra de carbono al 30 %.

La red neuronal artificial (ANN) está inspirada en el sistema nervioso biológico y se utiliza para resolver una amplia variedad de problemas científicos y de ingeniería complejos. Esta técnica computacional es especialmente útil para la simulación de cualquier correlación que sea difícil de describir con un modelo físico, ya que tienen la capacidad de aprender con el ejemplo y reconocer patrones en una serie de valores de entrada y salida de casos de ejemplo. Esta notable capacidad de modelado es útil en el estudio de problemas complicados, que normalmente no pueden resolverse con las teorías físicas existentes u otros enfoques matemáticos. El enfoque de red neuronal se ha introducido recientemente en el campo de estudio del comportamiento al desgaste de polímeros y compuestos. Velten [26] y Zhang [27] estuvieron entre los primeros pioneros en explorar este enfoque de los compuestos poliméricos y utilizaron modelos neuronales para predecir el volumen de desgaste de los termoplásticos reforzados con fibras cortas/partículas. Velten investigó el potencial de las técnicas redes neuronales para predecir y analizar el comportamiento de desgaste de compuestos PA46 reforzados con fibras cortas (fibras de vidrio y carbono, PTFE y grafito). A partir del análisis, señalaron que los resultados de la predicción del desgaste basada en redes neuronales presentaban datos muy viables y

prometedores para fines de diseño de materiales, estudios de parámetros sistemáticos y análisis de propiedades de los compuestos poliméricos. Zhang y Friedrich [17] revisaron la aplicación de varios principios del enfoque de redes neuronales para predecir la vida útil a la fatiga, el rendimiento frente al desgaste y las propiedades mecánicas dinámicas del compuesto de matriz polimérica. Zhang y Barkoula [28] predijeron la tasa de desgaste específica y el coeficiente de fricción del compuesto de poliamida reforzado con fibras cortas utilizando una red neuronal. Se han desarrollado varias investigaciones donde se presentan nuevos modelos y su eficiencia ha sido comparada con las redes neuronales para problemas específicos. Alimam y Hinnawi [29] desarrollaron modelos basados en redes neuronales y un modelo ANFIS (Adaptative Neuro-Based Fuzzy Inference System) para predecir el desgaste abrasivo de una lámina de aleación de aluminio 3105. Finalmente, concluyeron que, la predicción del modelo neuronal y ANFIS mostraron un buen desempeño en comparación con los datos experimentales. Gopan y Arun [30] presentaron un enfoque de red neuronal integrando algoritmo genético (GA) basado en la evolución para el modelado predictivo y la optimización de los parámetros de mecanizado. En [31], [32], [33], [34] se implementó un sistema de red neuronal artificial para modelar el comportamiento del desgaste y predecir los datos de vida de herramientas de mecanizado. La predicción del desgaste por parte del modelo neuronal proporciona una alternativa a los métodos de prueba actuales que consumen mucho tiempo y dinero. Desde entonces, esta técnica se ha utilizado con éxito en el campo de la tribología y la ingeniería, que incluye el desgaste de compuestos polimetilmetacrilato [35], índice de calidad de agua [7], rendimiento de diferentes dispositivos de energía solar (SE) [36], pérdidas primarias en régimen de flujo turbulento [37]. Dhande y Phate [38] describieron una metodología de superficie de respuesta y el modelado matemático basado en redes neuronales, ambos modelos determinaron el volumen perdido provocado por el deslizamiento en seco de un pasador de PTFE, y se compararon utilizando métodos estadísticos. El rendimiento de la red se ha corroborado en varias investigaciones utilizando el coeficiente de determinación, el error cuadrático medio, el error relativo medio, el error cuadrático medio y la regresión como medida de calidad.

2.2. Fundamentación teórica

2.2.1. Tribología

La tribología es la rama de la ciencia y la tecnología que se encarga de la interacción de superficies en movimiento relativo con materias asociadas [17]. Es considerada una nueva ciencia que estudia el análisis de la fricción, el desgaste y la lubricación de los materiales. Estos tres procesos se afectan mutuamente con causas y efectos que interactúan entre ellos [18]. El avance de la industria da origen a la búsqueda de reducir o controlar la fricción y el desgaste de los materiales, prolongar su vida útil, mejorar su eficiencia, encontrar nuevas fuentes de energía, etc. En la Figura 1 se observan dos superficies rugosas colocadas en contacto. La fricción entre ambas provoca la rugosidad y se crean puntos de contacto discreto. Los sistemas mecánicos en los que las superficies no se deslizan o ruedan entre sí son poco comunes. Los avances en tribología han respaldado gran parte del progreso de la ingeniería, la tecnología moderna, el transporte, la generación de energía, la ingeniería médica, la ciencia de los alimentos y la cosmética. La fricción juega un papel central en el desempeño de muchos sistemas mecánicos. La baja fricción es deseable e incluso esencial en varias circunstancias. El funcionamiento satisfactorio de las articulaciones, bisagras de puertas, articulaciones de la cadera humana o soportes de puentes, exigen una fuerza de fricción baja [10]. El trabajo realizado para superar la fricción en los cojinetes y engranajes de las máquinas se disipa en forma de calor y se logra un aumento en la eficiencia general. La baja fricción no es necesariamente beneficiosa en todos los casos. En frenos y embragues, la fricción adecuada y controlada es fundamental para disipar la energía cinética y transferir el par.

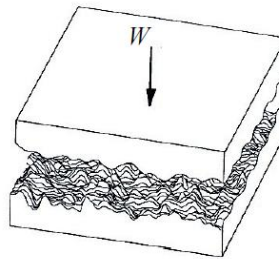


Figura 1: Fricción producida por dos superficies rugosas colocadas en contacto. W corresponde al peso de la superficie superior. Fuente: [19]

El desgaste es producido cuando dos superficies sólidas se deslizan una sobre la otra [39]. En la Figura 2 se presentan los diferentes mecanismos de desgaste entre dos cuerpos: desgaste por adhesión, desgaste por abrasión, desgaste por fatiga y desgaste por corrosión. La distinción entre desgaste abrasivo y desgaste por adhesión es artificial, ya que en algunas condiciones se puede generar residuos, que luego provocan un mayor desgaste por abrasión [40], [41].

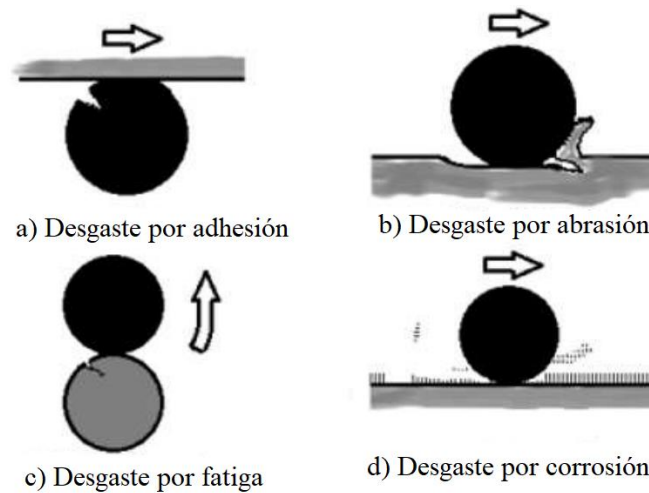


Figura 2: Mecanismos de desgaste. En el panel a) se muestra el desgaste por adhesión entre dos superficies deslizantes. En el panel b) se presenta el desgaste por abrasión entre una superficie dura y una más blanda. En el panel c) se muestra el desgaste por fatiga entre dos materiales por cargas cíclicas. En el panel d) se observa el desgaste por corrosión causado por la formación de una película en la superficie del material. Fuente: [42]

El desgaste abrasivo ocurre cuando una superficie o partículas duras pasan sobre una superficie blanda, causando pérdida de material [39]. En la figura 3 se presentan las situaciones de desgaste abrasivo más comunes. En el caso (a) la superficie más dura de las dos que rozan (abrasión de un cuerpo sobre otro) produce abrasión en la superficie más blanda. En el caso (b), la parte dura es un tercer cuerpo, generalmente partículas abrasivas atrapadas entre las dos superficies, lo suficientemente duras como para producir abrasión en una o ambas superficies del contacto (abrasión por tercer cuerpo) [40].

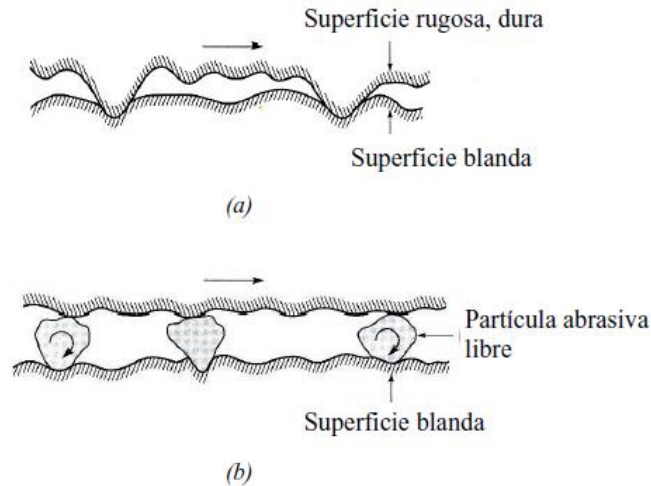


Figura 3: Formas de desgaste abrasivo. En el caso (a) se muestra el desgaste por rugosidad o picos de una superficie directamente sobre otra superficie blanda. En el caso (b) se presenta el desgaste entre dos superficies en contacto con partículas abrasivas libres. Fuente: [19]

A medida que dos superficies se desgastan una contra la otra, hay tres cuestiones clave que deben tenerse en cuenta:

- Las tensiones (tanto normales como de tracción) provocan daños mecánicos en los materiales, normalmente por deformación plástica o por fractura; este daño puede ocurrir en el nivel de aspereza, o en un nivel mucho más macroscópico.
- El trabajo de fricción producido cuando las superficies se deslizan unas contra otras, se disipa en forma de calor, lo que resulta en aumentos de temperatura en los cuerpos. Los cambios macroscópicos de temperatura pueden influir en las propiedades del material (como el límite elástico).
- Las superficies de todos los materiales interactúan con el entorno en el que se encuentran, por adsorción de átomos o moléculas o por reacción química. Para los metales en el aire, esta reacción generalmente implica la formación de una capa delgada de óxido metálico en la superficie [21].

2.2.2. Machine Learning (ML)

En 1936 se considera el punto de partida de la Inteligencia Artificial (IA) [26]. Alan Turing aplicó sus teorías para demostrar que una máquina era capaz de ejecutar un proceso cognitivo, y que podría dividirse en múltiples pasos individuales. El término IA propuesto por el programador John McCarthy fue acuñado en 1956. En 1966 se creó el primer programa informático capaz de comunicarse con los humanos. En 1972 la IA entró en el campo de la medicina ayudando a los médicos en el diagnóstico y el tratamiento. 'Deep Blue', una computadora de IBM venció al campeón mundial de ajedrez en 1997 y en 2005 un robot de Stanford condujo de forma autónoma 131 millas [6]. La inteligencia artificial podría definirse como un algoritmo informático que realiza un trabajo inteligente [43]. Ejemplos cotidianos de IA son: predicciones impulsadas por IA de Google (Google Maps), aplicaciones de viajes compartidos (Uber) [44]. El Machine Learning (ML) se define como el análisis de algoritmos informáticos capaces de aprender automáticamente en base a un entrenamiento adecuado [45], [46].

Tabla 1: Diferencias entre la Inteligencia Artificial y Machine Learning

Inteligencia artificial	Machine Learning
La inteligencia artificial es la teoría y el desarrollo de sistemas informáticos capaces de realizar tareas que normalmente requieren inteligencia humana.	El aprendizaje automático es una rama de la inteligencia artificial y se trata de la construcción y el estudio de sistemas que pueden aprender de los datos.
El objetivo es simular la inteligencia natural para resolver un problema complejo.	El objetivo es aprender de los datos de una determinada tarea para maximizar el rendimiento de la máquina.

2.2.3. Red neuronal artificial (RNA).

El concepto de Redes neuronales artificiales (RNA) se introduce desde el tema de la biología. La red neuronal biológica juega un papel importante en el cuerpo humano [47]. Una red neuronal es una red de millones y millones neuronas interconectadas entre sí. Las neuronas interconectadas realizan un procesamiento paralelo dentro del cuerpo humano. La Figura 4 corresponde a una neurona biológica que procesa información de una neurona a otra con la ayuda de algún cambio eléctrico y químico. Está compuesto por un cuerpo

celular y dos tipos de ramas que se extienden en forma de árbol: el axón y las dendritas. El proceso de recepción y envío de señales es a través de las dendritas. La neurona envía señales en picos de actividad eléctrica a través de un soporte largo y delgado conocido como axón el cual divide estas señales a través de la sinapsis y las envía a las demás neuronas [6].

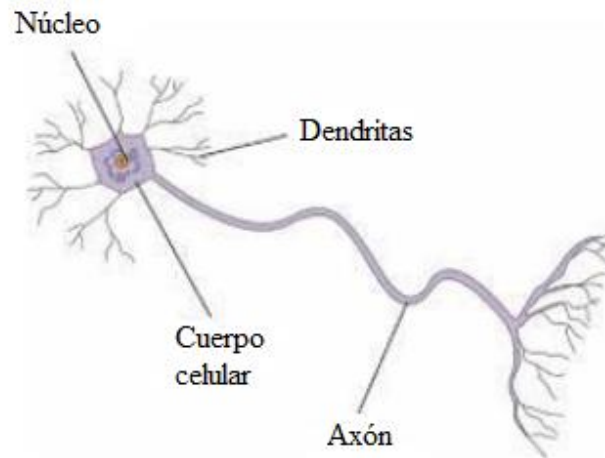


Figura 4: Estructura de una red neuronal biológica conformado por dendritas que transmiten estímulos al núcleo, donde se procesa la información y se envía una respuesta a través del axón. Fuente: [6]

Las redes neuronales son un paradigma computacional con una gran cantidad de unidades de cómputo simple (neuronas). Los modelos neuronales están interconectados en red para realizar tareas de procesamiento de datos de forma paralela. Las redes neuronales exhiben características excelentes como: procesamiento de información de alta velocidad, capacidades de mapeo, tolerancia a fallas, adaptabilidad, generalización y robustez. Las redes neuronales son una herramienta poderosa e inteligente para modelar, predecir y optimizar el desempeño de diferentes sistemas de ingeniería. Por lo tanto, se ha aplicado a problemas complejos de ingeniería no lineal en diferentes aplicaciones del mundo real con una reducción significativa de costo y tiempo [48].

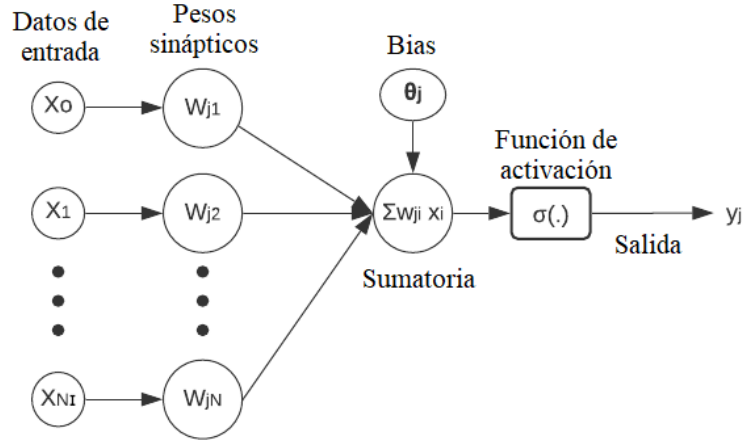


Figura 5: Modelo neuronal artificial conformado por la capa de entrada, capas ocultas y una capa de salida. Cada nodo se conecta con otro y tiene asociado un peso y un umbral, finalmente la capa de salida proporciona la predicción final.

Una red neuronal está compuesta por un conjunto de neuronas artificiales emulando conexiones sinápticas. Cada neurona artificial se representa como una unidad compuesta de conexiones de entrada, un núcleo central de proceso y una salida. La Figura 5 representa un modelo neuronal compuesto de una serie de entradas $x_i (i = [1, \dots, N_I])$ que equivalen a las dendritas, las cuales reciben la estimulación. La fuerza de conexión entre una entrada y una neurona se representa por el valor del peso w_{ji} . Los pesos son modificados en respuesta al entrenamiento que se le aplique a la red, dependiendo de la arquitectura específica o de sus reglas de aprendizaje. Los valores de peso negativos reflejan conexiones inhibitorias, mientras que los valores positivos designan conexiones excitadoras, si w_{ji} es cero, significaría que no existe una comunicación directa entre las neuronas i y j . La suma ponderada de las entradas y los pesos sinápticos llega a la función de activación $\sigma(x)$ que se encarga de calcular el nivel o estado de activación de la neurona y_j en función de los datos de entrada [27]. Un rango aceptable de salida suele estar entre 0 y 1, o -1 y 1. Existen diferentes tipos de funciones, desde simples funciones de umbral a funciones no lineales como la tangente hiperbólica o sigmoidea. La función de activación puede ser modificada adicionándole un término θ_j , denominado umbral (threshold o bias), el cual tiene la función de sumar o restar potencial postsináptico, a esto se le conoce como grado de libertad adicional de una neurona [49]. Una neurona artificial j se describe mediante la relación

$$y_j = \sigma \left(\sum_{i=1}^{N_I} w_{ji} x_i \pm \theta_j \right). \quad (1)$$

Los algoritmos de aprendizaje automático de las redes neuronales emplean una variedad de métodos estadísticos, probabilísticos y de optimización para aprender de la experiencia y detectar patrones útiles a partir de conjuntos de datos grandes, no estructurados y complejos [50]. El entrenamiento es un proceso de optimización donde se ajustan los parámetros internos del modelo neuronal (pesos y umbrales) para adaptarse a los datos de entrenamiento. Al inicio del entrenamiento los valores de w_{ji} generalmente se establecen de forma aleatoria o como valores nulos. En el proceso de entrenamiento, los mismos se modifican hasta encontrar una respuesta aceptable para la tarea deseada. La variación de pesos sinápticos se realiza en función de la regla de aprendizaje, la cual determina el error de la red. Existen dos métodos para el entrenamiento de la red neuronal: aprendizaje supervisado y aprendizaje no supervisado.

El aprendizaje supervisado se realiza mediante un entrenamiento controlado. La respuesta de la red se determina a partir de los valores de entrada [48]. En este aprendizaje, se proporcionan tanto las entradas como las salidas de la red. La red neuronal deberá procesar los valores resultantes y compararlos con las salidas deseadas. En la Figura 6 se representa el sistema de ajuste de pesos provocado por la corrección de errores en la red [51]. Este proceso ocurre una y otra vez a medida que los pesos se ajustan continuamente. Sin embargo, existe la posibilidad que algunas redes nunca aprendan. Este problema es debido a que los datos de entrada no contienen la información específica de la que se deriva la salida deseada. Las redes no convergen si no existen los suficientes datos para permitir un aprendizaje completo. El entrenamiento supervisado retiene un conjunto de datos que se utilizarán para determinar si el sistema ha generalizado la respuesta, o simplemente memorizó sus datos de entrenamiento de manera no significativa. [52].

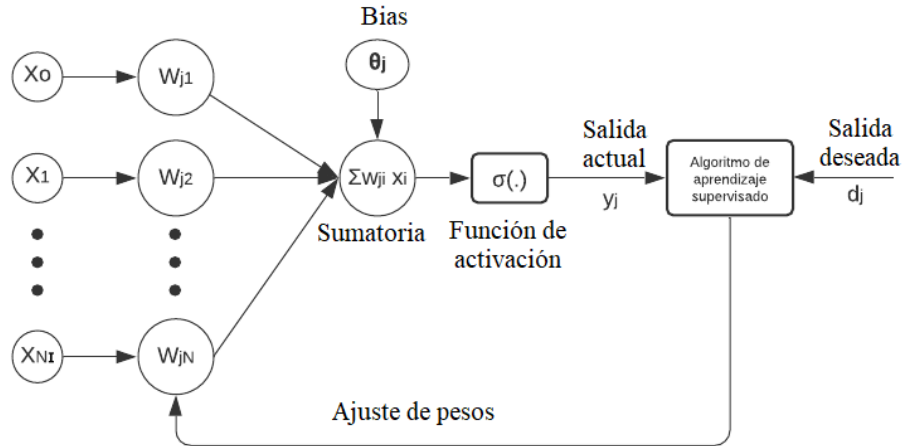


Figura 6: Modelo neuronal de aprendizaje supervisado basado en el reajuste de pesos de conexión de cada neurona controlado por un agente externo, y determina la respuesta de salida de la red.

En el aprendizaje no supervisado o auto-organizativo se le proporciona las entradas, pero no las salidas deseadas [53]. El sistema decidirá las funciones más adecuadas para agrupar los datos de entrada. La mayor parte de trabajos de investigación en base a redes neuronales se realizan en sistemas con aprendizaje supervisado. Este método utiliza la regla de aprendizaje competitivo, en el cual, las neuronas compiten entre sí para activarse frente a un dato de entrada. La red opera en el modo denominado "winner-takes-all", donde la neurona que gana la competencia se enciende o es premiada, y las demás neuronas son desactivadas [22]. La técnica de backpropagation o propagación hacia atrás se utiliza para implementar la retroalimentación adaptativa y ajustar los pesos durante el entrenamiento [54].

2.2.4. Modelo de Perceptrón Multicapa.

El perceptrón multicapa (MLP de su nombre en inglés Multilayer Perceptron) es de tipo unidireccional y se basa en un aprendizaje supervisado [43]. Las MLP son las redes neuronales de retroalimentación más utilizadas debido a su operación rápida, facilidad de implementación y requisitos de conjuntos de entrenamiento más pequeños. En la Figura 7 se presenta la estructura de un MLP con $k = 1$, el cual consta de tres capas secuenciales: capas de entrada, ocultas y de salida. La capa oculta procesa y transmite la información

de entrada a la capa de salida. Un MLP con un número insuficiente o excesivo de neuronas en la capa oculta causa problemas de mala generalización y sobreajuste [5]. No existe un método analítico para determinar el número de neuronas en la capa oculta. Por lo tanto, el investigador recurrirá al procedimiento de ensayo y error. La operación de un MLP con una capa oculta y una capa de salida lineal se puede expresar a través de la relación

$$y_k = \sum_{j=1}^{N_H} w_{kj} \sigma \left(\sum_{i=1}^{N_I} w_{ji} x_i \pm \theta_j \right) \pm \theta_k, \quad (2)$$

donde x_i representa cada una de las N_I neuronas de entradas de la red, N_H es el número de neuronas en la capa oculta, y_k a la k -ésima salida de la capa final, w_{ji} son los pesos de la capa oculta y θ_j sus umbrales, w_{kj} los pesos de la capa de salida y θ_k sus umbrales.

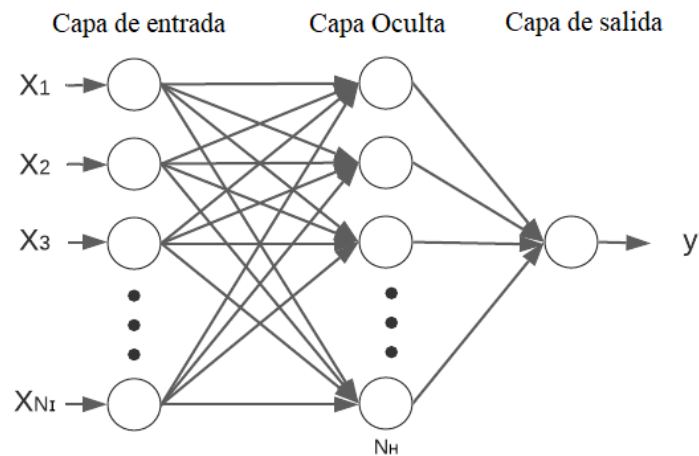


Figura 7: Modelo neuronal Perceptrón Multi-Capa compuesto por la capa de entrada, capa oculta y una capa de salida con $k = 1$.

La Figura 7 representa una red feedforward donde las señales neuronales se propagan hacia delante a través de las capas de la red. Estas redes no presentan conexiones hacia atrás y tampoco son auto-recurrentes, ni laterales [55]. Este tipo de redes son especialmente útiles en aplicaciones de reconocimiento o clasificación de patrones [49]. La construcción de una red neuronal debe considerar el tipo de modelo y arquitectura dependiendo del problema de estudio. El programador define los hiperparámetros del modelo, pero no los parámetros del modelo [56].

- Hiperparámetros del modelo: Son parámetros ajustables que permiten controlar el proceso de entrenamiento de un modelo neuronal.
- Parámetros del modelo: son las variables que se aprenden durante la fase de entrenamiento con los conjuntos de datos.

Los hiperparámetros del modelo son los siguientes:

- Número de capas ocultas.
- Número de neuronas en cada capa.
- Función de activación.
- Optimizador.
- Función de pérdida.
- Tasa de aprendizaje o Learning rate (LR).
- Batch size.

Parámetros del modelo:

- Pesos: Determinan la influencia entre la entrada de una neurona y su salida.
- Sesgos: son entradas adicionales que se utilizan para garantizar que incluso cuando las entradas son cero, habrá una activación en la neurona.

2.2.5. Función de activación

La función de activación está adjunta a cada neurona y determina la salida del modelo. Esta función debe ser computacionalmente eficiente, ya que es utilizada en repetidas ocasiones en la fase de entrenamiento. La red neuronal no debe presentar diferentes funciones de activación, ya que esto provoca diferentes comportamientos en la salida de cada neurona, provocando inconvenientes en el resultado final [25].

Una función escalón está basada en umbrales, es decir, se determina si activar o no una neurona. Una vez se define el umbral, si la entrada de la neurona es mayor a dicho umbral, la neurona se activa y su salida es igual a la entrada, es decir se puede expresar como

$$\sigma(x) = \begin{cases} 1 & \text{si } x \geq 0; \\ 0 & \text{si } x < 0. \end{cases} \quad (3)$$

La Figura 8 representa la función escalón con umbral igual a 0. La limitación más significativa de esta función es la imposibilidad de gestionar la salida de varios valores.

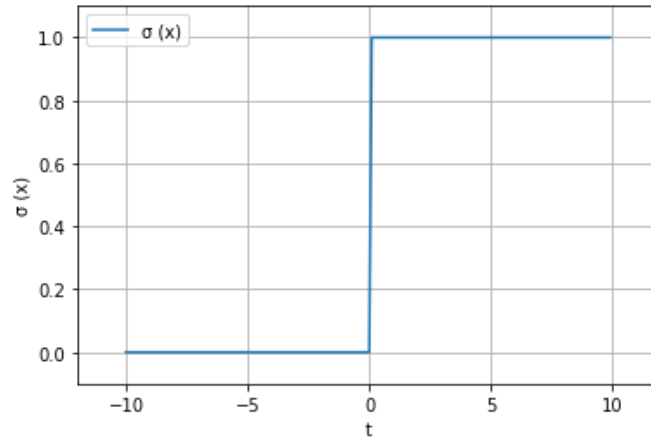


Figura 8: Gráfica de la función escalón con un umbral igual a 0.

Una función de activación lineal genera una salida proporcional a la entrada. Por lo tanto, la salida es directamente proporcional a la multiplicación del peso y la entrada

$$\sigma(x) = x. \quad (4)$$

La Figura 9 representa la función lineal como función de activación.

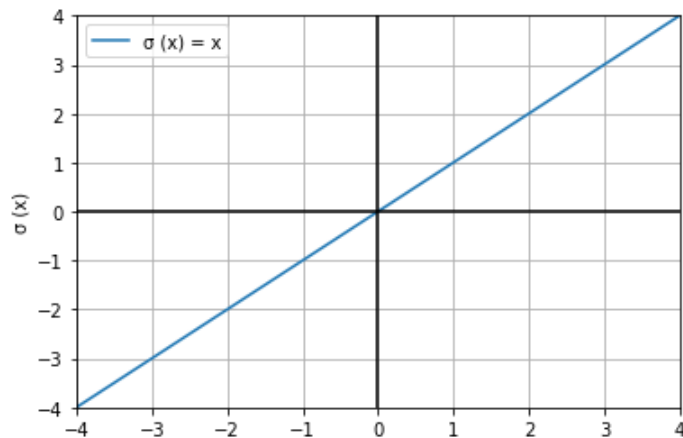


Figura 9: Gráfica de la función lineal como función de activación.

La función Sigmoidal es utilizada en las redes neuronales de tipo MLP, debido a la flexibilidad y rango de resultados que esta ofrece. Esta función permite aprender variaciones no lineales. La función sigmoide es una función de variable real diferenciable. En este caso, se utiliza la función logística:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (5)$$

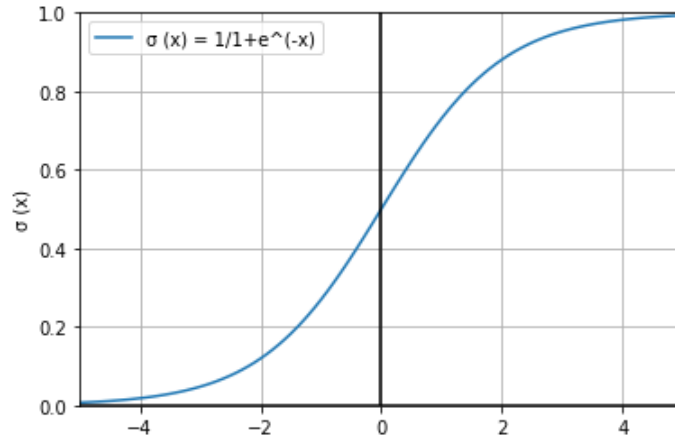


Figura 10: Gráfica de la función sigmoidea como función de activación.

En la Figura 10 se observan los valores en los que se aplica σ según el valor de x . Esta función se satura cuando su argumento es muy positivo o negativo, es decir, la función tiende a una curva muy plana e insensible a pequeños cambios en su entrada. Esta función presenta las siguientes características:

- Es una función no lineal.
- Su salida se encuentra en el rango (0 y 1) es suave, continua y diferenciable. La salida puede interpretarse como un valor de probabilidad.
- Es computacionalmente más costoso ya que involucra exponenciales.

La función de activación de tangente hiperbólica se define como la relación entre el seno hiperbólico y coseno hiperbólico:

$$\sigma(x) = \tanh(x) = \frac{\sinh(x)}{\cosh(x)}. \quad (6)$$

La Figura 11 representa la función de activación \tanh . Esta función es una versión matemáticamente transformada y desplazada de la función sigmoidea. Además, tiene un rango de valores de salida entre -1 y 1 .

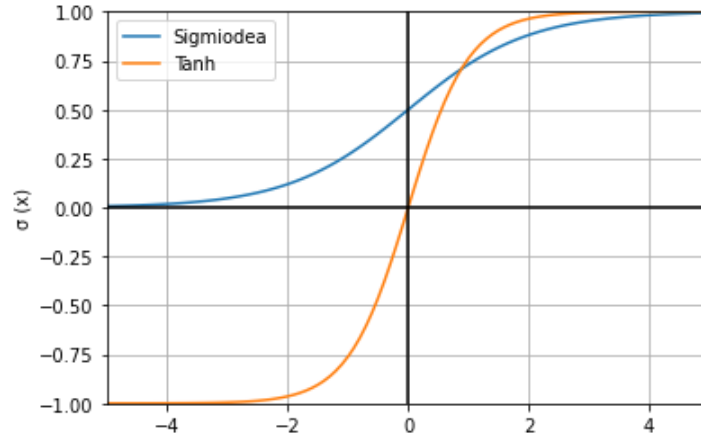


Figura 11: Gráfica de la función de activación de la tangente hiperbólica representada por la línea naranja y la función Sigmoidea de color azul.

La función de activación ReLU es lineal para todos los valores de entrada positivos y cero para todos los valores negativos, se expresa matemáticamente como:

$$\sigma(x) = \begin{cases} x & \text{si } x \geq 0; \\ 0 & \text{si } x < 0. \end{cases} \quad (7)$$

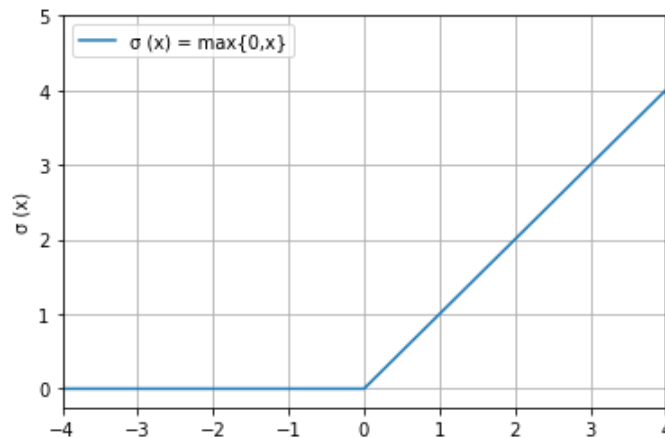


Figura 12: Gráfica de la función de activación de una ReLU con valores positivos a partir de cero, se caracteriza por no presentar valores negativos.

ReLU se considera una función de activación lineal por partes. La Figura 12 muestra el valor que toma la ReLU según el valor de x . Una ventaja de la función ReLU es su bajo costo computacional. Uno de los principales inconvenientes de tener valores de cero es el llamado “dying ReLU”. Una neurona ReLU está “desactivada” si se encuentra en el lado negativo. Cuando una neurona se transforma en negativa, es poco probable que se active nuevamente.

La función de activación Leaky ReLU es una variante de ReLU diseñada para solucionar el problema de “dying ReLU”. Leaky ReLU transforma los valores entrantes multiplicando los negativos por un coeficiente rectificativo. Por lo general, este valor se establece en 0,01. Leaky ReLU tiene una pequeña pendiente para valores negativos, en lugar de cero. Este sistema de ecuaciones se puede expresar matemáticamente a través de la expresión

$$\sigma(x) = \begin{cases} x & \text{si } x \geq 0; \\ 0.01x & \text{si } x < 0. \end{cases} \quad (8)$$

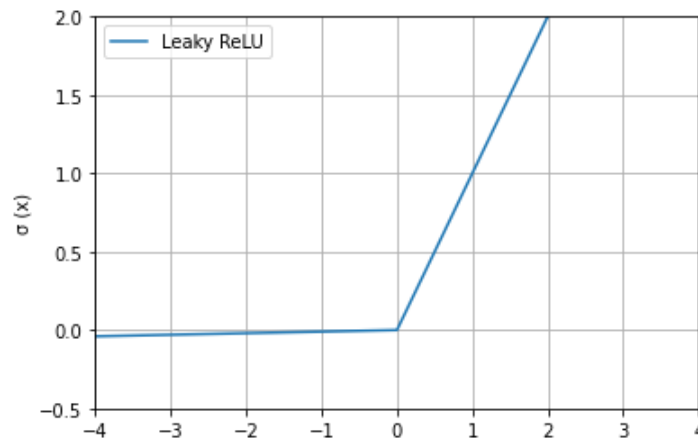


Figura 13: Gráfica de la función de activación Leaky ReLU con una pendiente para los valores negativos, evitando que las neuronas se desactiven.

Se muestra en la Figura 13 el valor que toma la Leaky ReLU según el valor de x .

2.2.6. Backpropagation

El algoritmo de backpropagation es un método de optimización por el descenso del gradiente (GD) de una función de error. Se busca minimizar el error cuadrático medio (MSE) de la salida calculada por la red neuronal [57]. Su diagrama se presenta en la Figura 14 donde se involucra tres fases: la fase del cómputo hacia adelante con los patrones de entrenamiento de entrada, el cálculo del error cometido y la propagación hacia atrás del error con el ajuste de los pesos (y umbrales) del modelo [58]. La combinación de pesos minimiza la función de error y genera una solución del problema de aprendizaje [59]. Este método requiere el cálculo del gradiente de la función de error en cada paso de iteración. Se debe garantizar la continuidad y diferenciabilidad de la función de error. Una red

neuronal de backpropagation es un gráfico computacional cuyos nodos son unidades de computación que transmiten información numérica de nodo a nodo. Cada unidad de cálculo es capaz de evaluar una única función primitiva de su entrada [60].

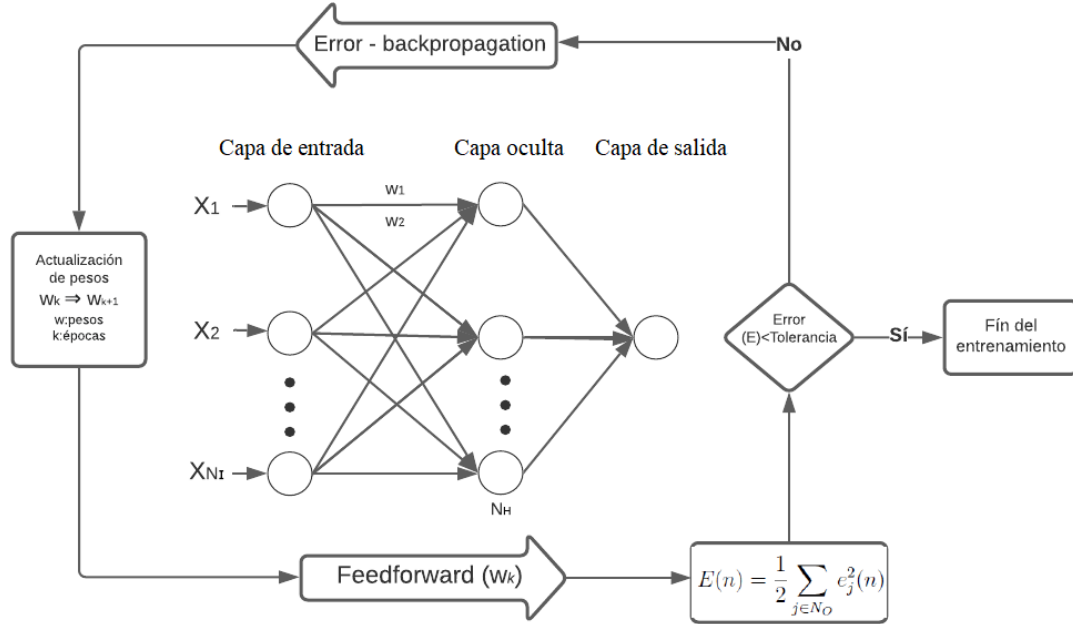


Figura 14: Diagrama del algoritmo de backpropagation. La primera fase consiste en el Feedforward el cual envía las señales hacia delante, los resultados se controlan a través del error de tolerancia. Si los resultados no presentan un valor satisfactorio se aplica el backpropagation, que consiste en reajustar los pesos y bias de las neuronas de la red.

El procedimiento de backpropagation consiste en actualizar los pesos y umbrales en base al error cuadrático medio. La red actualiza los parámetros para minimizar el error de todas las neuronas de entrenamiento. La red trabaja con aprendizaje supervisado y necesita un conjunto de datos de entrenamiento que describan cada entrada y su valor de salida esperado o target. Se dispone de un conjunto de N pares de entrenamiento que se denota $[x_N, t_N]$ donde x es la entrada y t la salida deseada [49]. El error de la neurona de salida j en la iteración o época de entrenamiento n se define como:

$$e_j(n) = t_j(n) - y_j(n). \quad (9)$$

El valor del error cuadrático para la neurona j se establece como $\frac{1}{2} e_j^2(n)$. Consecuentemente se define el valor instantáneo $E(n)$ de la suma los errores cuadráticos

de $\frac{1}{2}e_j^2(n)$ en todas las neuronas de la capa de salida N_o . La suma instantánea del error cuadrático de la red es escrita de la siguiente manera:

$$E(n) = \frac{1}{2} \sum_{j \in N_o} e_j^2(n), \quad (10)$$

donde el conjunto N_o incluye a todas las neuronas de la capa de salida de la red. El error cuadrático medio se determina como la suma anterior normalizada con respecto al tamaño N del conjunto de entrenamiento

$$MSE = \frac{1}{N} \sum_{n=1}^N \frac{1}{2} \sum_{j \in N_o} e_j^2(n). \quad (11)$$

La suma instantánea de los errores cuadráticos $E(n)$, y el error cuadrático medio MSE , son función de todos los parámetros libres de la red (pesos y umbrales). Para el conjunto de entrenamiento N , MSE representa la función de coste como medida del proceso de aprendizaje. El objetivo del aprendizaje es ajustar los parámetros libres de forma que se minimice el MSE . Para minimizar el error se los pesos deben ser actualizados patrón a patrón [27]. La corrección de los pesos sinápticos se realiza mediante el incremento en estos pesos con la relación $w'_{ji}(n) = w_{ji}(n) + \Delta w_{ji}(n)$. El algoritmo aplica una corrección $\Delta w_{ji}(n)$ al peso sináptico $w_{ji}(n)$, el que es proporcional al gradiente $\frac{\partial E(n)}{\partial w_{ji}(n)}$.

De acuerdo con la regla de la cadena, es posible expresar este gradiente como:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \cdot \frac{\partial e_j(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial u_j(n)} \cdot \frac{\partial u_j(n)}{\partial w_{ji}(n)}. \quad (12)$$

Donde $y_j(n) = \sigma(u_j(n))$ y $u_j(n) = \sum_{i=1}^{N_I} w_{ji}(n)x_i(n)$. Esto resulta en:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n)\sigma'(u_j(n))x_i(n). \quad (13)$$

Entonces, la corrección $\Delta w_{ji}(n)$ aplicado a $w_{ji}(n)$ es definido por la regla delta:

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)}, \quad (14)$$

donde η es una constante que determina la tasa o velocidad de aprendizaje. El uso del signo menos denota el descenso por el gradiente del error en el espacio de los pesos.

Uniendo $\Delta w_{ji}(n)$ y $\frac{\partial E(n)}{\partial w_{ji}(n)}$ se obtiene:

$$\Delta w_{ji}(n) = \eta e_j(n) \sigma' (u_j(n)) x_i(n). \quad (15)$$

El gradiente local para una neurona j es igual al producto de la correspondiente señal de error $e_j(n)$ y la derivada de la función de activación de las neuronas asociadas a la capa a la cual pertenece dicha neurona. Se puede identificar dos casos distintos dependiendo de la localización de la neurona j de la red. El primer caso, la neurona j está en la capa de salida. Para calcular la señal del error en cada nodo se debe restar de la salida deseada la señal que se obtuvo naturalmente. En el segundo caso, la neurona j está ubicada en una capa oculta por lo tanto no existe una respuesta deseada específica para esta neurona [27]. Esto provoca que se determine la señal de error recursivamente para todas las neuronas asociadas a la neurona. Se define el gradiente local de una neurona oculta j como:

$$\Delta w_{ji}(n) = \eta \frac{\partial E(n)}{\partial w_{ji}(n)} \sigma' (u_j(n)) x_i(n), \quad (16)$$

donde la neurona j es una neurona oculta, y

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k(n) \frac{\partial e_k(n)}{\partial y_k(n)}. \quad (17)$$

El error de una neurona de salida k está definido como $e_k(n) = t_k(n) - y_k(n)$, por lo tanto $y_k(n) = \sigma(u_k(n))$ y $u_k(n) = \sum_{j=1} w_{kj}(n) y_j(n)$. Mediante la regla de la cadena para la derivada $\frac{\partial e_k(n)}{\partial y_j(n)}$ se obtiene la ecuación:

$$\frac{\partial E(n)}{\partial y_j(n)} = - \sum_k e_k(n) \sigma'(u_k(n)) w_{kj}(n). \quad (18)$$

Se obtiene la regla de actualización de pesos en la capa oculta como:

$$\Delta w_{ji}(n) = \eta \sum_k e_k(n) \sigma'(u_k(n)) w_{kj}(n) \sigma'(u_j(n)) x_i(n). \quad (19)$$

2.2.7. Algoritmos de los Optimizadores

Los métodos de optimización se utilizan para lograr una solución óptima de problemas no lineales en aplicaciones de aprendizaje profundo [61]. Los métodos de optimización Stochastic Gradient Descent (SGD), Adadelta, Adam, RmsProp, Nadam son los más utilizados en la biblioteca Keras. Los métodos de optimización permiten una actualización rápida de los parámetros dentro del algoritmo. Los algoritmos de optimización del descenso de gradiente se pueden clasificar en tres categorías: descenso de gradiente por lotes o batch, descenso de gradiente estocástico (SGD) y descenso de gradiente por mini-batch. SGD es el algoritmo de optimización de descenso de gradientes más utilizado en los últimos años. El proceso de convergencia de SGD no es lo suficientemente estable y puede generar sobreajuste conduciendo a un desempeño insatisfactorio en múltiples épocas [27]. El algoritmo de optimización del descenso del gradiente de Adagrad varía la tasa de aprendizaje para adaptarse durante todo el proceso de convergencia. Este enfoque funciona bien en conjuntos de datos dispersos. Adagrad presenta una debilidad, la tasa de aprendizaje se ralentiza durante el entrenamiento. El rendimiento del aprendizaje es insatisfactorio cuando se presentan múltiples épocas. El algoritmo RMSprop divide la tasa de aprendizaje por un promedio de gradientes cuadrados y decae exponencialmente, resolviendo el problema de la tasa de aprendizaje decreciente. El algoritmo de optimización Adadelta también resuelve el problema de la disminución de la tasa de aprendizaje, similar a RMSprop. El algoritmo de optimización del descenso del gradiente de Adam combina las ventajas de RMSprop junto con el método de Momentum. Este algoritmo presenta un buen desempeño en muchas tareas de aprendizaje profundo [27].

El descenso de gradientes (GD) es un algoritmo de optimización utilizado para entrenar un modelo de aprendizaje automático. Es un algoritmo para encontrar un mínimo local de una función diferenciable. Modifica los valores de los parámetros de una función de forma iterativa para minimizar una función de pérdida a su mínimo local [62]. Dicha derivada tiene información que ayuda a minimizar o maximizar esa función, ya que es la pendiente de la función $f(x)$ en el punto x . El mínimo local es un punto donde $f(x)$ es menor que todos los puntos vecinos [27]. La definición matemática es la siguiente: x^* es un mínimo local si $f(x^*) \leq f(x)$ para todo x en X a distancia $\epsilon > 0$ de x^* , donde X es un espacio de la función de pérdida. Un máximo local es un punto donde $f(x)$ es más alto que f en todos los puntos vecinos. Los puntos donde la derivada es cero se denominan puntos estacionarios o puntos críticos. El mínimo y el máximo globales son los puntos más bajos y altos en todo el dominio. Un problema relacionado con la optimización en el aprendizaje automático se debe a que las funciones objetivas pueden tener más de un mínimo local. Cuando se utilizan funciones con múltiples entradas derivadas parciales, $\frac{\delta}{\delta x_i} f(x)$ representar la derivada de f con respecto a x_i [46].

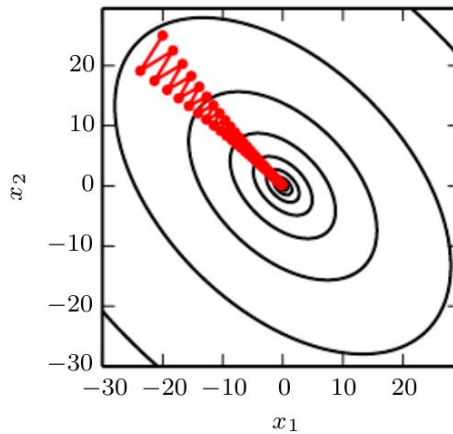


Figura 15: Método de descenso de gradiente utilizado para minimizar una función de pérdida y converger hacia el valor mínimo de una función mediante un proceso iterativo.

Fuente: [46]

El gradiente $\nabla_x f(x)$ es el vector que contiene todas las derivadas parciales de la función f con respecto a cada variable x_i . Mediante múltiples variables es posible disminuir f siguiendo la dirección dada por gradiente negativo:

$$x' = x - \epsilon \nabla_x f(x), \quad (20)$$

dónde ϵ es la tasa de aprendizaje, x son los parámetros del modelo (pesos y sesgos), f es la función objetivo, $\nabla_x f(x)$ es el gradiente de la función objetivo y x' son los parámetros del modelo ajustados [46]. La Figura 15 se representa el descenso de gradiente.

El descenso de gradiente estocástico (SGD) es el algoritmo optimizador más común utilizado en el aprendizaje profundo. SGD calcula el gradiente utilizando una pequeña parte aleatoria de los datos y se actualiza después de cada muestra [27]. La optimización utilizada por SGD se realiza en una pequeña parte de los datos denominada mini-batch, $B = \{x^{(1)}, \dots, x^{(m')}\}$. El mini-batch suele ser una pequeña parte de todo el conjunto de datos que debe ser lo suficientemente pequeño para ser computacionalmente eficiente y lo suficientemente grande para que la función de costo del mini-batch represente correctamente la función de costo real de todo el conjunto de datos de entrenamiento. Se define la función de pérdida como $L(x_i, y_i, \theta)$, donde θ representa los parámetros del modelo, lo cual podría introducirse como:

$$g = \frac{1}{m'} \nabla_{\theta} \sum_{i=1}^{m'} L(x_i, y_i, \theta), \quad (21)$$

Y el paso del SGD se establecería como:

$$\theta' = \theta - \epsilon g. \quad (22)$$

La tasa de aprendizaje (LR) es un parámetro importante en SGD y es recomendable disminuir dicho valor en la fase de entrenamiento [47]. Cuando la tasa de aprendizaje es demasiado grande, la fase de aprendizaje muestra oscilaciones violentas y puede conducir a un mínimo local, en cambio, si es demasiado pequeña, la fase de aprendizaje podría atascarse con un valor de función de alto costo. La propiedad más importante de SGD es que el tiempo de cálculo requerido por cada actualización no aumenta con el tamaño del conjunto de datos [46].

El método de momentum está diseñado para acelerar el proceso de aprendizaje. Acumula un promedio móvil decreciente exponencial de los gradientes pasados y continúa moviéndose en esas direcciones. Este método acumula un promedio móvil decreciente exponencial de los gradientes anteriores y continúa desplazándose en esas direcciones. Un promedio decreciente exponencial es un promedio móvil exponencial particular (Exponential Moving Average, EMA) [47]. La nueva ley de actualización se representa como:

$$v = \alpha v' - \epsilon \frac{1}{m} \nabla_{\theta} \sum_{i=1}^{m'} L(x_i, y_i, \theta) ; \quad (23)$$

$$\theta = \theta' + v. \quad (24)$$

El parámetro $\alpha \in [0,1]$ y controla las caídas exponenciales. Los parámetros v se denominan velocidad y acumula los gradientes anteriores.

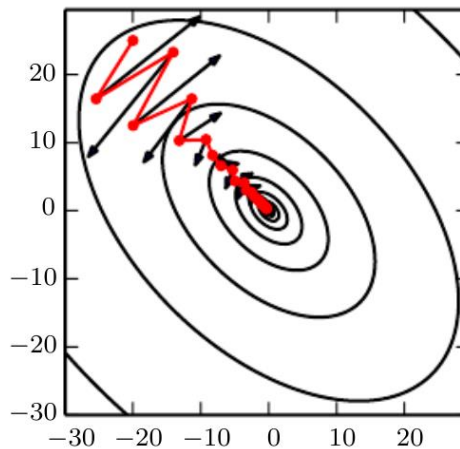


Figura 16: Método de momentum frente a gradiente descendiente. La línea roja representa el algoritmo de momentum diseñado para acelerar el proceso de optimización mediante impulsos, y las flechas negras al gradiente descendiente basado en la información del gradiente anterior hasta encontrar el mínimo local. Fuente: [46]

La Figura 16 representa las principales diferencias entre el método de Momentum y el descenso de gradiente. La línea roja corresponde al Momentum, las flechas negras representan el paso del descenso de gradiente en ese punto. Los pasos del gradiente

desperdician tiempo computacional moviéndose hacia adelante y hacia atrás. El algoritmo de momentum de desplaza correctamente hasta la convergencia. En el descenso de gradiente, el tamaño del paso se alinea con el gradiente del último aprendizaje, mientras que con el momentum el paso depende del tamaño y la dirección de la actualización anterior.

El algoritmo de RMSProp se caracteriza por una tasa de aprendizaje adaptativo. La tasa de aprendizaje aumenta cuando el signo de las derivadas permanece igual, si cambian de signo, la tasa de aprendizaje disminuye [6]. RMSProp adapta individualmente las tasas de aprendizaje de todos los parámetros del modelo escalando inversamente a la raíz cuadrada de la suma de todos sus valores históricos al cuadrado. Esto representa una disminución rápida de la tasa de aprendizaje para aquellos parámetros con valores grandes y una disminución lenta para parámetros con valores pequeños [46]. Es necesario introducir algunos parámetros para analizar el algoritmo RMSProp: ϵ es la tasa de aprendizaje, ρ es la tasa de caída de la media móvil ponderado exponencial y δ es una constante pequeña para evitar la división por cero (generalmente se establece en 10^{-6}). A partir de la ecuación del gradiente, RMSProp continúa con los gradientes cuadrados de la siguiente manera:

$$r = \rho r' + (1 - \rho)g \odot g ; \quad (25)$$

$$\Delta\theta = \frac{-\epsilon}{\sqrt{\delta + r}} \odot g ; \quad (26)$$

$$\theta = \theta' + \Delta\theta . \quad (27)$$

El algoritmo Adam es una combinación del método RMSProp y Momentum. El algoritmo introduce la media móvil exponencial y los gradientes al cuadrado. Introduce correcciones basadas en los gradientes pasados (método de Momentum) y correcciones basadas en la raíz cuadrada de los gradientes pasados cuadrados (método RMSProp). Incorpora una corrección de sesgo para ambos términos de impulso que es importante en el primer paso donde el promedio móvil en declive se lleva a cero ya que ρ_1 y ρ_2 están cerca de cero. Se introducen nuevas definiciones: ρ_1 y ρ_2 en $[0,1)$ son las tasas de desintegración exponencial para el primer y segundo momentum (los números sugeridos son 0,9 y 0,999),

δ es una constante pequeña para evitar la división por cero, s es el primer momentum (gradientes), r es la variable de segundo momentum (gradientes cuadrados). El planteamiento matemático surge la ecuación del gradiente:

$$s = \rho_1 s' + (1 - \rho_1)g ; \quad (28)$$

$$r = \rho_1 r' + (1 - \rho_2)g \odot g ; \quad (29)$$

$$\hat{s} = \frac{s}{1 - \rho_1} ; \quad (30)$$

$$\hat{r} = \frac{r}{1 - \rho_2} ; \quad (31)$$

$$\Delta\theta = -\epsilon \frac{\hat{s}}{\sqrt{\hat{r} + \delta}} ; \quad (32)$$

$$\theta = \theta' + \Delta\theta . \quad (33)$$

2.2.8. Función de pérdida

El entrenamiento de un modelo neuronal se evalúa mediante la función de pérdida. Son funciones que estiman el grado de inconsistencia entre el valor predicho de su modelo y el valor verdadero. En el proceso de validación, la función de pérdida es la encargada de maximizar o minimizar los errores [46].

El error cuadrático medio o MSE evalúa la calidad de un predictor [13]. MSE proporciona información sobre las diferencias entre los valores objetivo y predichos de la siguiente manera:

$$MSE = \frac{1}{N} \sum_{j=1}^N (t_j - y_j)^2 , \quad (34)$$

dónde N representa el número de muestras, y_j es el j^{th} valor predicho y t_j es el j^{th} valor objetivo. Los valores más bajos de MSE indican un mejor ajuste de la RNA.

El error absoluto medio o MAE determina el promedio de la diferencia absoluta entre los valores verdaderos y predichos. MAE y MSE se han utilizado para evaluar el rendimiento del modelo durante muchos años, no existe consenso sobre la métrica más adecuada para

determinar los errores de un modelo neuronal [34]. MAE está definido como:

$$MAE = \frac{1}{N} \sum_{j=1}^N |t_j - y_j|. \quad (35)$$

El coeficiente de determinación o R^2 refleja la bondad del ajuste de un modelo a la variable objetivo [31]. Se utiliza con fines descriptivos y mide la capacidad explicativa del modelo estimado. R^2 se representa matemáticamente como:

$$R^2 = 1 - \frac{\sum(t_j - y_j)^2}{\sum(t_j - \bar{y}_j)^2}, \quad (36)$$

donde \bar{y}_j es la media de los valores predichos. R^2 tiene una escala muy intuitiva, ya que su rango presenta valores entre 0 a 1. Si el resultado es cercano a 1, representa un mayor ajuste del modelo a la variable esperada. Cero representa que el modelo propuesto no se ajusta a la variable esperada.

CAPITULO III

MARCO METODOLÓGICO

3.1. Ubicación.

Las placas de látex y partículas de caucho reciclado se elaboraron en la empresa Color Plast, ubicada en la ciudad de Ambato. Los ensayos de desgaste abrasivo se realizaron en los laboratorios de la Facultad de Ingeniería Civil y Mecánica de la Universidad Técnica de Ambato.

3.2. Equipos y materiales.

Software:

- Lenguaje de programación Python 3.9.6.
- Google Colaboratory.
- Librerías TensorFlow y Keras.
- Autodesk Fusion 360

Materiales:

- Látex de policloruro de vinilo (PVC)
- Gránulo de caucho reciclado (0 - 0,841 mm)
- Horno industrial para vulcanizado.
- Abrasímetro Taber.
- Máquina CNC (Control numérico por computadora).

3.3. Tipo de investigación.

3.3.1. Investigación bibliográfica.

La investigación se realizó mediante la revisión de material bibliográfico referente al tema de estudio tales como revistas científicas, artículos científicos, libros electrónicos, tesis doctorales, y páginas web.

3.3.2. Investigación aplicada.

La investigación aplicada se enfatiza en generar conocimiento para solventar problemas actuales de la sociedad. Se basa principalmente en buscar nuevas alternativas tecnológicos, combinando de manera simultánea la teoría y los resultados. El presente proyecto presenta la importancia de fabricar un nuevo material como alternativa al caucho convencional.

3.3.3. Investigación cuantitativa.

La presente investigación consiste en recolectar y analizar datos numéricos obtenidos a través de ensayos realizados en laboratorio. Se presentan los valores obtenidos, predicciones, relacionar las variables independientes con las dependientes, y la obtención de resultados generales. Mediante un análisis estadístico se logrará probar la hipótesis establecida.

3.4. Prueba de hipótesis.

- **Hipótesis Alternativa**

Aplicando un modelo de red neuronal artificial en un material compuesto de látex con partículas de caucho reciclado, se podrá predecir el desgaste abrasivo con un coeficiente de determinación, $R^2 > 90\%$.

- **Hipótesis nula**

Aplicando un modelo de red neuronal artificial en un material compuesto de látex con partículas de caucho reciclado, no se podrá predecir el desgaste abrasivo con un coeficiente de determinación, $R^2 > 90\%$.

3.5. Población o muestra.

La población objetivo del trabajo constituyen las placas elaboradas a partir del material compuesto de látex y gránulo de caucho reciclado, en sus diferentes configuraciones y variación en la composición porcentual de matriz – refuerzo. La Tabla 2 detalla las configuraciones de las placas analizadas en la investigación.

Tabla 2: Conjunto de datos

Composición volumétrica	Peso Aplicado (g)	Velocidad de giro (RPM)	Número de ciclos	Número de probetas
1. 60% Látex 40% Gránulo	250	60	500	27
	500	70	750	
	1000	80	100	
2. 70% Látex 30% Gránulo	250	60	500	27
	500	70	750	
	1000	80	100	
3. 80% Látex 20% Gránulo	250	60	500	27
	500	70	750	
	1000	80	100	
Total				81

3.6. Recolección de información.

La recolección de información se realizó a través de: artículos científicos, libros, revistas, investigaciones y tesis doctorales, obteniendo la información más relevante para el desarrollo de esta investigación, además, se recolectaron 81 datos experimentales aplicando diferentes variaciones volumétricas, utilizando 65 datos para el entrenamiento y 16 datos para la validación del modelo neuronal, equivalentes al 80% y 20% del total de datos, respectivamente. La Tabla 3 describe las preguntas y técnicas básicas para la recolección de información, es decir, proporciona los métodos adecuados para obtener la información necesaria que permita alcanzar los objetivos planteados en la investigación.

Tabla 3: Recolección de información

Preguntas básicas	Explicación
¿Para qué?	Obtención de datos de desgaste abrasivo para el entrenamiento y validación de la red neuronal.
¿De qué personas u objetos?	Placas de compuesto de látex con partículas de caucho reciclado.
¿Sobre qué aspectos?	Desgaste abrasivo
¿Quién, Quiénes?	Investigador
¿Cuándo?	Junio – septiembre 2021
¿Dónde?	La recolección de información se la obtuvo en el laboratorio de la facultad de Ingeniería Civil y Mecánica de la UTA.
¿Cuántas veces?	Se realizó 81 placas del material compuesto.
¿Qué técnicas de recolección?	Se utilizó el método de observación experimental.
¿Con qué?	Norma ISO 9352.
¿En qué situación?	Horario establecido por los laboratorios de la facultad de Ingeniería Civil y Mecánica de la UTA.

3.7. Procesamiento de la información y análisis estadístico.

La Figura 17 representa la esquematización gráfica del algoritmo utilizado para la obtención de las 81 placas mediante ensayos experimentales. La arquitectura del modelo neuronal se desarrolló en el software Tensorflow. Los 81 datos fueron analizados y entrenados de forma conjunta, los mismos que se dividieron de la siguiente manera: 65 datos para entrenamiento y 16 datos para validación. Los resultados fueron verificados aplicando estadísticos de prueba para aceptar o rechazar la hipótesis nula. El coeficiente de correlación obtenido fue superior al 95%, por lo tanto, se rechazó la hipótesis nula y se acepta la hipótesis alternativa, la cual sugiere que los cambios de composición volumétrica, carga aplicada, velocidad de desgaste y número de revoluciones afectan de manera directa al proceso de desgaste abrasivo.

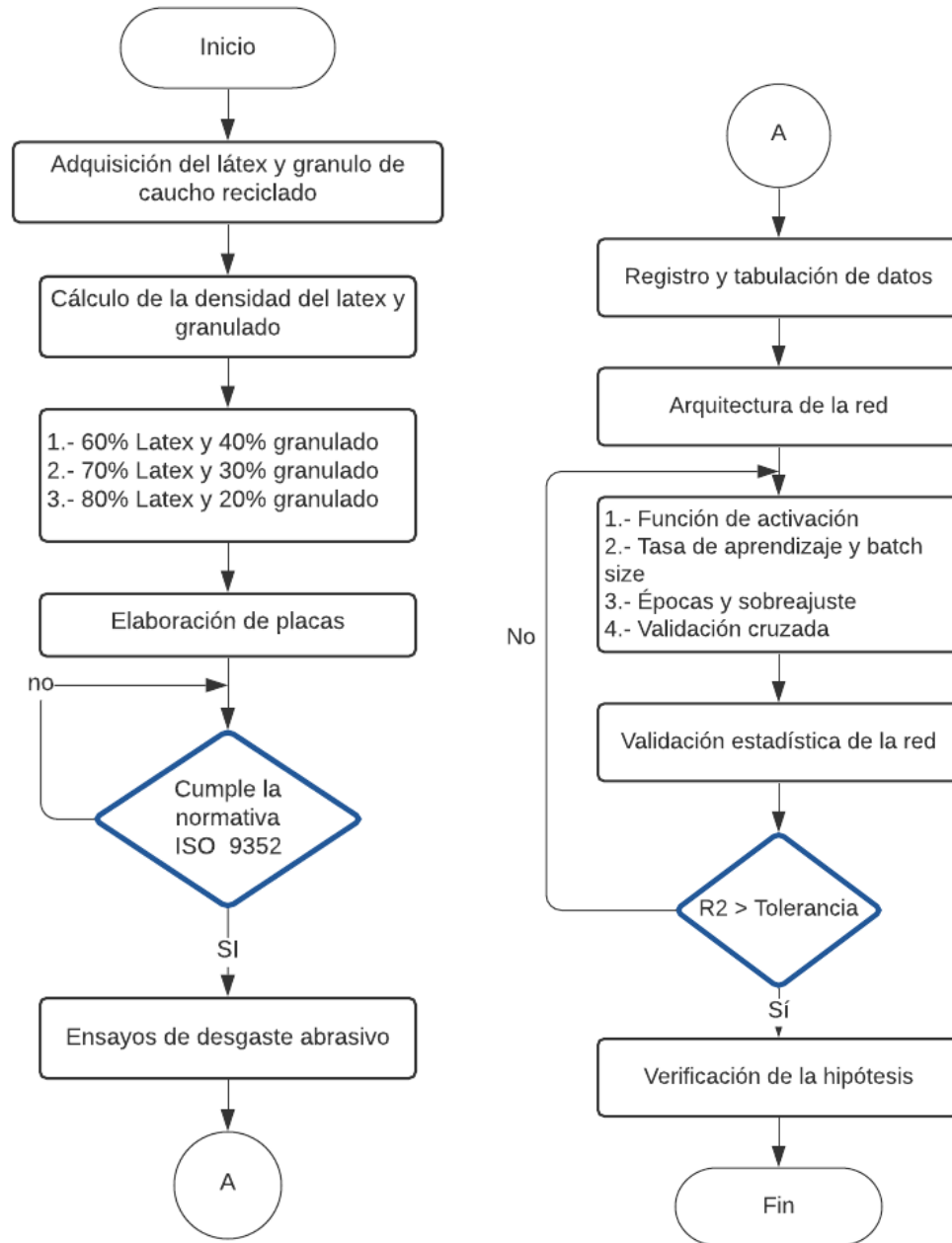


Figura 17: Diagrama del proceso metodológico de la investigación.

3.8. Variables de la investigación

La arquitectura neuronal se determina en base a los parámetros de entrada y salida establecidos para el entrenamiento. La correlación entre variables determinará su influencia en la predicción del desgaste abrasivo. El porcentaje volumétrico corresponde a la variable de entrada más relevante durante esta investigación. A partir de esta función

se definen las funciones de carga aplicada, velocidad de giro y número de revoluciones. De esta forma, la carga aplicada establecerá el nivel de fricción entre la muela abrasiva y la placa. Por su parte, la velocidad de giro y el número de revoluciones expresan el nivel de material desgastado por ciclo. Por último, el desgaste abrasivo corresponde a la variable dependiente, indicando el desprendimiento o pérdida de material en base a las variables independientes. La Tabla 4 describe un resumen de las variables dependientes e independientes utilizadas en esta investigación:

Tabla 4: Variables de estudio

Variables Independientes.	Variable Dependiente.
1. Porcentaje de material	1. Desgaste abrasivo
2. Carga aplicada	
3. Velocidad de giro	
4. Número de revoluciones	

CAPITULO IV

RESULTADOS Y DISCUSIÓN

El material compuesto se elaboró en base a la Norma ISO 9352 en la empresa “Color Plast”. La empresa analizó la propuesta y concluyó que el ensayo especificado en la norma estaba dentro del alcance de los equipos. Para la adquisición del material se tomó en cuenta la calidad, costos y logística en cuanto a transportación. El Látex Sintético se lo adquirió de la empresa “Color Plast”. El material provisto por esta empresa cumplió con los requisitos de calidad. El gránulo de caucho reciclado se adquirió en la empresa “PRONEUMACOSA”, ubicada en la provincia de Cotopaxi. La empresa se dedica al reciclaje de neumáticos mediante un proceso ecológico y libre de contaminación. En el presente estudio se optó por la granulometría más pequeña ofrecida por la empresa, la cual fue del tipo: 0 – 1,2 mm de acuerdo con el tamiz #18.

El material empleado para fabricar el molde fue el aluminio, ya que presenta varias características y ventajas en comparación a otros materiales.

- **Peso:** El aluminio es, en peso una tercera parte del acero. La ligereza de este material facilita la manipulación de las placas.
- **Mecanizado:** En comparación con el acero el aluminio es un material más fácil de mecanizar.
- **Moldes para la producción de piezas de plástico:** El aluminio es un gran conductor del calor y disipa de 3 a 4 veces más rápido.

El diseño del molde se realizó en el Software Autodesk Fusion 360. Las dimensiones de las probetas se determinaron en base a la norma ISO 9352. La figura 18 presenta las dimensiones de las probetas: 111 x 6,35 mm de diámetro y espesor respectivamente, perforación central de ¼ de pulgada.

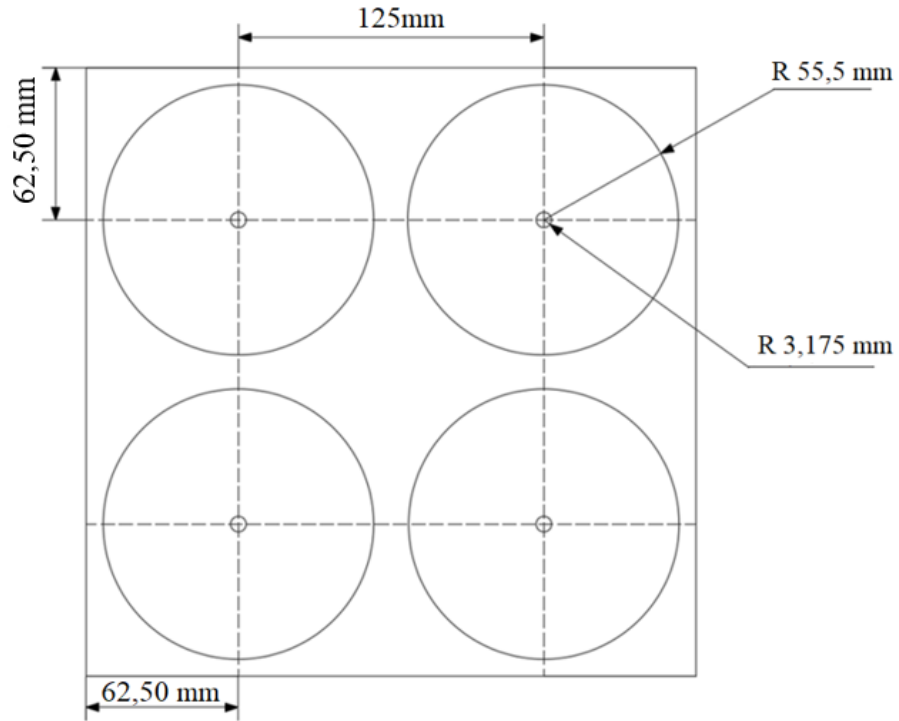


Figura 18: Diseño CAD del molde de las probetas de acuerdo con la norma ISO 9352, conformado por 4 cavidades homogéneas.

El maquinado del molde se realizó en la empresa “Color Plast” con la ayuda de la máquina CNC. La Figura 19 muestra el acabado superficial realizado mediante lija de grano fino.

Tabla 5: Dimensiones del molde

Tipo de molde	Dimensiones de placa de aluminio. (mm)	Dimensiones de las cavidades. (mm)	Número de cavidades.	Volumen parcial por mecanizar. (cm ³)	Volumen total por mecanizar. (cm ³)	Duración de mecanizado. (horas)
Molde para ensayos de desgaste abrasivo	250*250*8	Diámetro: 111 Espesor: 6,35	4	244,988	979,954	6

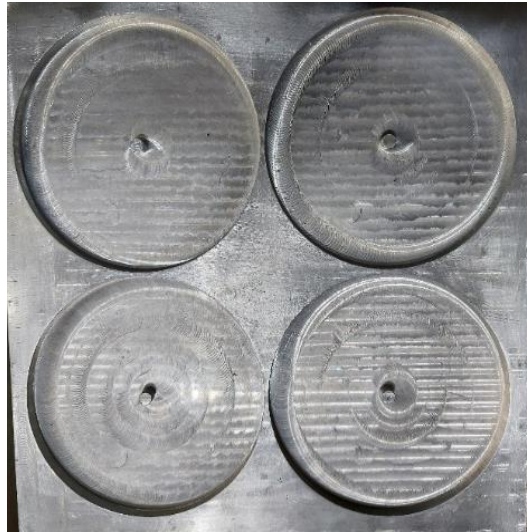


Figura 19: Molde mecanizado en la maquina CNC de la empresa "ColorPlast".

El volumen de las placas individuales se obtuvo a partir de la relación entre la altura y la diferencia de radios:

$$V = \pi \times h \times (R^2 - r^2) \quad (37)$$

$$V = \pi \times 6,35mm \times [(111mm)^2 - (6,35mm)^2]$$

$$V = 244\,988,627\,mm^3$$

$$V = 244,988\,cm^3,$$

donde:

V = Volumen de cada placa.

h = Espesor del molde.

R = Radio mayor.

r = Radio menor.

El volumen de cada sección del molde fue de $244,988cm^3$. El molde cuenta con cuatro cavidades lo que representa un volumen total de $979,954cm^3$.

4.1. Densidad y volumen del látex sintético

La masa y el volumen del látex sintético se obtuvo a través de pruebas experimentales a temperatura ambiente. La masa del látex se determinó utilizando una probeta maciza elaborada por medio del proceso de vulcanizado. El pesaje de la probeta se realizó un total

de cinco veces. Los resultados se promediaron y se obtuvo un valor de 16,082 g.

Tabla 6: Mediciones de la masa de látex sintético.

Probeta de látex	Masa (g)
Ensayo 1	16,11
Ensayo 2	16,08
Ensayo 3	16,09
Ensayo 4	16,08
Ensayo 5	16,10
Promedio	16,082

El volumen del látex sintético se determinó con la ayuda de un vaso de precipitación graduada y 60 ml de agua. Se introdujo la probeta de látex dentro del vaso de precipitación junto con el agua. El volumen de la probeta de látex se determinó a partir de la diferencia de volúmenes (volumen final – volumen inicial), dicho proceso se realizó cinco veces. El promedio de la diferencia de volúmenes fue de 10,6 ml = 10,6 cm³.

Tabla 7: Mediciones del volumen del látex sintético

Probeta de látex	Volumen inicial (ml)	Volumen final (ml)	Diferencia de volúmenes (ml)
Ensayo 1	200	210	10
Ensayo 2	200	211	11
Ensayo 3	200	210	10
Ensayo 4	200	211	11
Ensayo 5	200	211	11
Promedio			10,6

La densidad del látex se representa como el cociente entre la masa y el volumen del látex

$$\rho_L = \frac{m_L}{V_L} \quad (38)$$

$$\rho_L = \frac{16,082g}{10,6cm^3}$$

$$\rho_L = 1,517 \frac{g}{cm^3},$$

donde:

ρ_L = Densidad del látex

m_L = Masa de látex

V_L = Volumen del látex

4.2. Densidad y volumen del gránulo de caucho reciclado

El tamaño del gránulo de caucho utilizado para las mediciones fue de 0 – 1,2 mm del tamiz #18. La masa del gránulo de caucho se determinó en base a una pequeña cantidad medida en la balanza electrónica. La medición se la realizó en cinco ocasiones, obteniendo un promedio de masa de 10 g. El volumen del granulado de caucho se determinó utilizando un vaso de precipitación graduado. Se introdujo 100 ml de látex sintético junto con el granulado de caucho. Se mezcló de forma continua ambas sustancias hasta conseguir una masa homogénea. El volumen del gránulo de caucho se obtuvo mediante la diferencia entre el volumen final e inicial. El proceso se replicó en cinco ocasiones, obteniendo un promedio de 11,2 ml equivalente a 11,2 cm³.

Tabla 8: Mediciones del volumen del granulado de caucho

Probeta de látex	Masa de PGCR (g)	Volumen inicial (ml)	Volumen final (ml)	Diferencia de volúmenes (ml)
Ensayo 1	10	100	110	10
Ensayo 2	10	100	111	11
Ensayo 3	10	100	112	12
Ensayo 4	10	100	112	12
Ensayo 5	10	100	111	11
Promedio				11,2

La densidad del granulado de caucho se representa como el cociente entre la masa y el volumen del granulado

$$\rho_g = \frac{m_g}{V_g}$$
$$\rho_g = \frac{10g}{11,2cm^3}$$
$$\rho_g = 0,892 \frac{g}{cm^3},$$

donde:

ρ_g = Densidad del granulado de caucho

m_g = Masa del granulado de caucho

V_g = Volumen del granulado de caucho

El cálculo de la fracción volumétrica de cada caso de estudio se estableció en base a la densidad individual de los materiales.

Tabla 9: Densidad de matriz y refuerzo

Descripción	Valor
Densidad del látex	1,517 g/cm ³
Densidad del granulado de caucho	0,892 g/cm ³
Volumen de cada probeta	244,988 cm ³

Composición volumétrica: 60% matriz (látex sintético) y 40% refuerzo (Gránulo de caucho reciclado). El volumen de gránulo y látex se determinó a través de:

$$V_L = 0,6 \times 244,988 \text{ cm}^3$$

$$V_L = 146,993 \text{ cm}^3$$

$$V_g = 0,4 \times 244,988 \text{ cm}^3$$

$$V_g = 97,995 \text{ cm}^3.$$

La masa del látex sintético y gránulo de caucho se estableció despejando la masa de la ecuación de la densidad

$$m_L = \rho_L \times V_L$$

$$m_L = 1,517 \frac{\text{g}}{\text{cm}^3} \times 146,993 \text{ cm}^3$$

$$m_L = 222,988 \text{ g}$$

$$m_g = \rho_g \times V_g$$

$$m_g = 0,892 \frac{\text{g}}{\text{cm}^3} \times 97,995 \text{ cm}^3$$

$$m_g = 87,411 \text{ g}.$$

Composición volumétrica: 70% matriz (látex sintético) y 30% refuerzo (Gránulo de caucho reciclado). El volumen de gránulo y látex se determinó a través de:

$$V_L = 0,7 \times 244,988 \text{ cm}^3$$

$$V_L = 171,492 \text{ cm}^3$$

$$V_g = 0,3 \times 244,988 \text{ cm}^3$$

$$V_g = 73,496 \text{ cm}^3 .$$

La masa del látex sintético y gránulo de caucho se estableció despejando la masa de la ecuación de la densidad

$$m_L = \rho_L \times V_L$$

$$m_L = 1,517 \frac{g}{cm^3} \times 171,492 cm^3$$

$$m_L = 260,153 g$$

$$m_g = \rho_g \times V_g$$

$$m_g = 0,892 \frac{g}{cm^3} \times 73,496 cm^3$$

$$m_g = 65,558 g .$$

Composición volumétrica: 80% matriz (látex sintético) y 20% refuerzo (Gránulo de caucho reciclado). El volumen de gránulo y látex se determinó a través de:

$$V_L = 0,8 \times 244,988 cm^3$$

$$V_L = 195,990 cm^3$$

$$V_g = 0,2 \times 244,988 cm^3$$

$$V_g = 48,997 cm^3 .$$

La masa del látex sintético y gránulo de caucho se estableció despejando la masa de la ecuación de la densidad

$$m_L = \rho_L \times V_L$$

$$m_L = 1,517 \frac{g}{cm^3} \times 195,99 cm^3$$

$$m_L = 297,318 g$$

$$m_g = \rho_g \times V_g$$

$$m_g = 0,892 \frac{g}{cm^3} \times 48,997 cm^3$$

$$m_g = 43,705 g .$$

La Tabla 10 presenta un resumen con los valores de volumen y masa para cada configuración volumétrica y la Tabla 11 presenta la denominación de cada tipo de composición volumétrica. La Tabla 12 detalla las diferentes configuraciones para la realización de las pruebas de desgaste abrasivo.

Tabla 10: Volumen y masa de las composiciones volumétricas

	Composición volumétrica	Volumen (cm³)	Masa (g)
1.	60% Látex	146,993	222,988
	40% Gránulo	97,995	87,411
2.	70% Látex	171,492	260,153
	30% Gránulo	73,496	65,558
3.	80% Látex	195,99	297,318
	20% Gránulo	48,997	43,705

Tabla 11: Denominación de las composiciones volumétricas

Denominación	Composición volumétrica
Tipo A	60% Látex 40% Gránulo
Tipo B	70% Látex 30% Gránulo
Tipo C	80% Látex 20% Gránulo

La elaboración del material compuesto se realizó mediante el proceso de estratificación manual. La Figura 20 representa la distribución homogénea de matriz y refuerzo, con una óptima impregnación, adherencia y compactación total de los elementos. La finalidad de esta mezcla homogénea es reducir la porosidad de las probetas, evitando la filtración de líquidos o humedad. El proceso de vulcanización requirió de un horno industrial calentado a 177 °C. El tiempo de vulcanizado aproximado es de tres minutos.

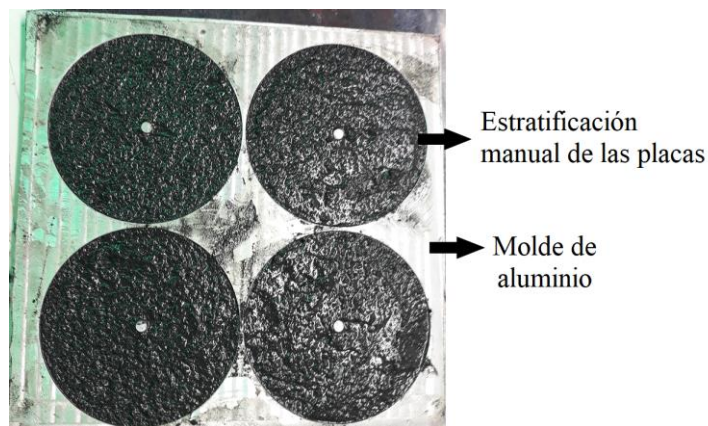


Figura 20: Proceso de estratificación manual de las probetas para encontrar una composición homogénea.

Tabla 12: Grupo de identificación de probetas para recolección de datos.

Número de probeta	Codificación	Composición volumétrica	Carga aplicada (g)	Velocidad de giro (RPM)	Número de ciclos
1	A1	60% L - 40% G	250	60	500
2	A2	60% L - 40% G	250	60	750
3	A3	60% L - 40% G	250	60	1000
4	A4	60% L - 40% G	250	70	500
5	A5	60% L - 40% G	250	70	750
6	A6	60% L - 40% G	250	70	1000
7	A7	60% L - 40% G	250	80	500
8	A8	60% L - 40% G	250	80	750
9	A9	60% L - 40% G	250	80	1000
10	A10	60% L - 40% G	500	60	500
11	A11	60% L - 40% G	500	60	750
12	A12	60% L - 40% G	500	60	1000
13	A13	60% L - 40% G	500	70	500
14	A14	60% L - 40% G	500	70	750
15	A15	60% L - 40% G	500	70	1000
16	A16	60% L - 40% G	500	80	500
17	A17	60% L - 40% G	500	80	750
18	A18	60% L - 40% G	500	80	1000
19	A19	60% L - 40% G	1000	60	500
20	A20	60% L - 40% G	1000	60	750
21	A21	60% L - 40% G	1000	60	1000
22	A22	60% L - 40% G	1000	70	500
23	A23	60% L - 40% G	1000	70	750
24	A24	60% L - 40% G	1000	70	1000
25	A25	60% L - 40% G	1000	80	500
26	A26	60% L - 40% G	1000	80	750
27	A27	60% L - 40% G	1000	80	1000
28	B1	70% L - 30% G	250	60	500
29	B2	70% L - 30% G	250	60	750
30	B3	70% L - 30% G	250	60	1000
31	B4	70% L - 30% G	250	70	500
32	B5	70% L - 30% G	250	70	750
33	B6	70% L - 30% G	250	70	1000
34	B7	70% L - 30% G	250	80	500
35	B8	70% L - 30% G	250	80	750
36	B9	70% L - 30% G	250	80	1000
37	B10	70% L - 30% G	500	60	500
38	B11	70% L - 30% G	500	60	750
39	B12	70% L - 30% G	500	60	1000
40	B13	70% L - 30% G	500	70	500

41	B14	70% L - 30% G	500	70	750
42	B15	70% L - 30% G	500	70	1000
43	B16	70% L - 30% G	500	80	500
44	B17	70% L - 30% G	500	80	750
45	B18	70% L - 30% G	500	80	1000
46	B19	70% L - 30% G	1000	60	500
47	B20	70% L - 30% G	1000	60	750
48	B21	70% L - 30% G	1000	60	1000
49	B22	70% L - 30% G	1000	70	500
50	B23	70% L - 30% G	1000	70	750
51	B24	70% L - 30% G	1000	70	1000
52	B25	70% L - 30% G	1000	80	500
53	B26	70% L - 30% G	1000	80	750
54	B27	70% L - 30% G	1000	80	1000
55	C1	80% L - 20% G	250	60	500
56	C2	80% L - 20% G	250	60	750
57	C3	80% L - 20% G	250	60	1000
58	C4	80% L - 20% G	250	70	500
59	C5	80% L - 20% G	250	70	750
60	C6	80% L - 20% G	250	70	1000
61	C7	80% L - 20% G	250	80	500
62	C8	80% L - 20% G	250	80	750
63	C9	80% L - 20% G	250	80	1000
64	C10	80% L - 20% G	500	60	500
65	C11	80% L - 20% G	500	60	750
66	C12	80% L - 20% G	500	60	1000
67	C13	80% L - 20% G	500	70	500
68	C14	80% L - 20% G	500	70	750
69	C15	80% L - 20% G	500	70	1000
70	C16	80% L - 20% G	500	80	500
71	C17	80% L - 20% G	500	80	750
72	C18	80% L - 20% G	500	80	1000
73	C19	80% L - 20% G	1000	60	500
74	C20	80% L - 20% G	1000	60	750
75	C21	80% L - 20% G	1000	60	1000
76	C22	80% L - 20% G	1000	70	500
77	C23	80% L - 20% G	1000	70	750
78	C24	80% L - 20% G	1000	70	1000
79	C25	80% L - 20% G	1000	80	500
80	C26	80% L - 20% G	1000	80	750
81	C27	80% L - 20% G	1000	80	1000

4.3. Ensayos de desgaste abrasivo

Los ensayos de desgaste abrasivo se realizaron en el Abrasímetro Taber. Las probetas fueron sometidas al desgaste por abrasión mediante el rozamiento de muelas durante un periodo de tiempo y una carga específica. Se observó el desprendimiento de micropartículas y las marcas de películas superficiales. El abrasímetro está incorporado a una aspiradora para absorber las partículas desprendidas. La eliminación de partículas evita la incidencia negativa durante la realización de los ensayos. El funcionamiento del abrasímetro está centrado en el giro del plato principal y los ejes de tipo horizontal, donde se colocan las muelas abrasivas. Estas giran en direcciones opuestas a la placa, produciendo el movimiento de desgaste requerido. Las muelas centran la abrasión en una zona específica la cual adopta la forma de un anillo provocando el desgaste en la zona exterior. La Figura 21 muestra el accionar de las muelas y la formación de la zona donde se produce el desgaste. La Tabla 13 presenta el procedimiento adecuado para la realización de las pruebas de desgaste abrasivo.

Tabla 13: Procedimiento para las pruebas de desgaste abrasivo.

Paso	Actividad	Descripción
1	Calibración	Se establece la velocidad y número de ciclos necesarios para la abrasión.
2	Colocación de las muelas	Se colocan las 2 muelas H-22 en los brazos del Abrasímetro
3	Fijación de la muestra a ser ensayada	Las probetas se ubican en el plato central, la misma que debe asegurarse mediante una tuerca.
4	Ubicación de las pesas	Se colocan las pesas en los brazos del Abrasímetro, dependiendo del ensayo a realizar.
5	Ensayo	Inicio del proceso de desgaste en base a los parámetros establecidos.
6	Extracción y Limpieza	Se retira la probeta y se retira las partículas sobrantes.
7	Pesaje	Se pesa la probeta en una balanza electrónica, registrando el peso antes y después del ensayo.
8	Rectificado de las muelas	Es aconsejable rectificar las muelas para mejorar mejores resultados en los próximos ensayos.



Figura 21: Proceso de desgaste abrasivo sobre una placa de ensayo.

El índice de desgaste Taber (IDT) es el indicador de la pérdida volumétrica de las probetas. El IDT determina el resultado de la rotura repetida de los enlaces de fricción entre la superficie de la placa en contacto con las muelas abrasivas. La fórmula del IDT se determinó entre la relación de la masa perdida por la placa y el número de ciclos realizados durante el ensayo

$$IDT = \frac{(P_1 - P_2) * F_c}{C}, \quad (39)$$

donde:

P_1 : Cantidad de masa previo del ensayo de abrasión (mg).

P_2 : Cantidad de masa posterior al ensayo de abrasión (mg).



C : Número de ciclos aplicados al ensayo.

F_c : Factor de corrección.

El factor de corrección se estableció en 100, debido a que el entrenamiento neuronal presentaba oscilaciones muy bruscas, evitando llegar a un valor de correlación adecuado. La función de pérdida presentaba valores elevados y R^2 inferiores a 0,9. Los resultados de los ensayos se registraron en las Tablas 14, 15 y 16.

4.1. Recolección de los resultados obtenidos del ensayo de desgaste abrasivo

Tabla 14: Recolección de datos con la configuración de 60 % Matriz – 40% PGCR

 UNIVERSIDAD TÉCNICA DE AMBATO FACULTAD DE INGENIERÍA CIVIL Y MECÁNICA CARRERA DE INGENIERÍA MECÁNICA 					
Datos Informativos					
Probetas: 27	Autorizada por:	Ing. Cristian Pérez		Realizado por:	Muñoz Pablo
Sitio de análisis y ensayo	Laboratorio de Materiales - FICM				
Tipo de estudio	Experimental			Fecha:	11/08/2021
Norma utilizada	ISO 9352				
Condiciones para el ensayo de desgaste abrasivo					
Características de la probeta				Características de la máquina	
Matriz	Látex	Fracción volumétrica	60%	Tipo de muela abrasiva	Calibrade H-22
Refuerzo	PGCR	Fracción volumétrica	40%		
Temperatura	19.7 °C			Número de muelas abrasivas	2
Dimensión	Ø 111 ± 0,3 mm				
Espesor	6,35 mm				
Resultados					
Codificación	Peso inicial (g)	Peso Final (g)	Índice de desgaste taber (IDT)	Porcentaje de pérdida de masa (%)	
A1	77,19	77,17	4	0,026	
A2	83,29	83,25	5,333	0,048	
A3	82,14	82,07	7	0,085	
A4	68,44	68,38	12	0,088	
A5	60,3	60,2	13,333	0,166	
A6	60,23	60,09	14	0,232	
A7	64,31	64,25	12	0,093	
A8	80,64	80,53	14,667	0,136	
A9	76,98	76,82	16	0,208	
A10	85,32	85,23	18	0,105	
A11	67,19	67,07	16	0,179	
A12	87,3	87,11	19	0,218	

A13	74,89	74,79	20	0,134
A14	80,69	80,52	22,667	0,211
A15	66,12	65,87	25	0,378
A16	65,72	65,59	26	0,198
A17	59,7	59,49	28	0,352
A18	65,88	65,58	30	0,455
A19	91,32	91,18	28	0,153
A20	66,27	66,05	29,333	0,332
A21	79,01	78,69	32	0,405
A22	80,15	80	30	0,187
A23	58,32	58,08	32	0,412
A24	57,27	56,92	35	0,611
A25	57,44	57,25	38	0,331
A26	81,5	81,19	41,333	0,380
A27	73,88	73,4	48	0,650

Fotografías posteriores al ensayo



Tabla 15: Recolección de datos con la configuración de 70 % Matriz – 30% PGCR

 UNIVERSIDAD TÉCNICA DE AMBATO FACULTAD DE INGENIERÍA CIVIL Y MECÁNICA CARRERA DE INGENIERÍA MECÁNICA 					
Datos Informativos					
Probetas: 27		Autorizada por: Ing. Cristian Pérez		Realizado por:	Muñoz Pablo
Sitio de análisis y ensayo		Laboratorio de Materiales - FICM			
Tipo de estudio		Experimental		Fecha:	11/08/2021
Norma utilizada		ISO 9352			
Condiciones para el ensayo de desgaste abrasivo					
Características de la probeta				Características de la máquina	
Matriz	Látex	Fracción volumétrica	70%	Tipo de muela abrasiva	Calibrade H-22
Refuerzo	PGCR	Fracción volumétrica	30%		
Temperatura		19.7 °C		Número de muelas abrasivas	2
Dimensión		Ø 111 ± 0,3 mm			
Espesor		6,35 mm			
Resultados					
Codificación	Peso inicial (g)	Peso Final (g)	Índice de desgaste taber (IDT)	Porcentaje de pérdida de masa (%)	
B1	64,25	64,18	22	0,109	
B2	56,12	55,98	24	0,249	
B3	54,77	54,55	25	0,402	
B4	65,91	65,8	26	0,167	
B5	70,38	70,2	28	0,256	
B6	60,05	59,8	30	0,416	
B7	49,4	49,27	30	0,263	
B8	61,83	61,62	34,667	0,340	
B9	76,27	75,97	36	0,393	
B10	78,55	78,4	36	0,191	
B11	76,04	75,78	37,333	0,342	
B12	68,04	67,68	39	0,529	
B13	71,44	71,26	38	0,252	
B14	64,9	64,62	40	0,431	

B15	58,86	58,47	43	0,663
B16	70,41	70,22	40	0,270
B17	65,84	65,54	42,667	0,456
B18	57,66	57,23	45	0,746
B19	73,43	73,23	44	0,272
B20	63,88	63,56	46,667	0,501
B21	70,5	70,05	48	0,638
B22	80,17	79,95	44	0,274
B23	71,7	71,35	46,667	0,488
B24	71,01	70,53	50	0,676
B25	64,2	63,98	22	0,343
B26	61,4	61,05	24	0,570
B27	63,37	62,87	25	0,789

Fotografías posteriores al ensayo



Tabla 16: Recolección de datos con la configuración de 70 % Matriz – 30% PGCR

		UNIVERSIDAD TÉCNICA DE AMBATO FACULTAD DE INGENIERÍA CIVIL Y MECÁNICA CARRERA DE INGENIERÍA MECÁNICA			
Datos Informativos					
Probetas: 27		Autorizada por:	Ing. Cristian Pérez	Realizado por:	Muñoz Pablo
Sitio de análisis y ensayo		Laboratorio de Materiales - FICM		Realizado por:	Muñoz Pablo
Tipo de estudio		Experimental		Fecha:	11/08/2021
Norma utilizada		ISO 9352			
Condiciones para el ensayo de desgaste abrasivo					
Características de la probeta				Características de la máquina	
Matriz	Látex	Fracción volumétrica	80%	Tipo de muela abrasiva	Calibrade H-22
Refuerzo	PGCR	Fracción volumétrica	20%	Tipo de muela abrasiva	Calibrade H-22
Temperatura		19.7 °C		Número de muelas abrasivas	2
Dimensión		$\varnothing 111 \pm 0,3$ mm			
Espesor		6,35 mm			
Resultados					
Codificación	Peso inicial (g)	Peso Final (g)	Índice de desgaste taber (IDT)	Porcentaje de pérdida de masa (%)	
C1	64,56	64,46	20	0,155	
C2	61,46	61,28	24	0,293	
C3	66,07	65,8	27	0,409	
C4	56,18	56,06	24	0,214	
C5	56,86	56,65	28	0,369	
C6	59,8	59,5	30	0,502	
C7	65,04	64,91	26	0,200	
C8	65,81	65,6	28	0,319	
C9	56,38	56,07	31	0,550	
C10	59,61	59,45	32	0,268	
C11	56,24	55,97	36	0,480	
C12	68,16	67,77	39	0,572	
C13	70,17	70	34	0,242	
C14	54,18	53,9	37,333	0,517	

C15	63,04	62,64	40	0,635
C16	61,75	61,55	40	0,324
C17	66,52	66,21	41,333	0,466
C18	61,83	61,38	45	0,728
C19	61,66	61,45	42	0,341
C20	61,7	61,37	44	0,535
C21	60,41	59,94	47	0,778
C22	61,25	60,99	52	0,424
C23	64,03	63,61	56	0,656
C24	65,05	64,45	60	0,922
C25	62,15	61,84	62	0,499
C26	60	59,49	68	0,850
C27	56,38	55,67	71	1,259

Fotografías posteriores al ensayo



4.4. Desarrollo de la Red neuronal artificial

El objetivo principal de esta investigación no es la predicción exacta del índice de desgaste abrasivo, sino demostrar un buen resultado con este enfoque. El objetivo es determinar un comportamiento particular para futuras aplicaciones en este campo. El enfoque de este proyecto está basado en TensorFlow. Esta es una plataforma Linux de 64 bits con interfaz de Python3. Es una biblioteca de software de código abierto desarrollada por Google y basada en un entorno de desarrollo integrado (IDE) para la programación. TensorFlow es el sistema de segunda generación de Google Brain y la primera versión se lanzó en febrero de 2017. La versión que se utiliza es la 2.7.0 disponible desde enero de 2019. Además, se utilizó la biblioteca Keras diseñada para el aprendizaje profundo. Keras es de código abierto y está escrito en lenguaje de programación Python. Se utilizaron TensorFlow 2.7.0 y Python 3.10.0 como lenguaje de programación. Se importaron las librerías más importantes para el entorno de programación, entre ellas se encuentra:

- Pandas: Biblioteca de software para Python y una herramienta para la manipulación y el análisis de datos.
- Scikit-learn: (abreviado Sklearn), es una biblioteca de software de código abierto para ML. El lenguaje de programación utilizado para Scikit-learn es principalmente Python.
- Numpy: Permite un trabajo sencillo con matrices, vectores y arrays.
- Matplotlib: Biblioteca de programación para visualizar scripts de Python.
- Keras: Biblioteca de aprendizaje profundo, compatible con TensorFlow.

Los códigos de programación que se desarrollaron fueron específicamente diseñados para este estudio. No se incluyeron códigos ya definidos. En primer lugar, es necesario subrayar dos aspectos importantes: la subdivisión de entrada/salida y la función de activación para la capa de salida. Las variables de entradas de la red son: variaciones volumétricas, la velocidad de giro, carga aplicada y el número de ciclos. La salida corresponde al índice de desgaste taber (IDT). En el caso de los modelos neuronales aplicados a la predicción, la función de activación más adecuada entre la capa oculta y la salida es la función de

activación lineal. El conjunto de datos obtenidas de los ensayos se organizó en un archivo .data con cinco columnas: Porcentaje, Peso, Velocidad, Ciclos e IDT. La columna de "Porcentaje" es categórica, por lo que fue necesario utilizar un "one-hot" para transformar dicha variable a numérica. La columna "Peso" corresponde a la carga aplicada para producir la abrasión y corresponde a los valores de 250, 500 y 1000 g, la columna "Velocidad" determina los valores de velocidad de abrasión, en este estudio fueron de 60, 70 y 80 RPM, y la columna "Ciclos" establece las revoluciones aplicadas al desgaste, las cuales se establecieron en 500, 700 y 1000 ciclos.

Tabla 17: Conjunto de datos de entrada

Variable de Entrada	Valores
Porcentaje	1. 60% Látex - 40% Gránulo
	2. 70% Látex - 30% Gránulo
	3. 80% Látex - 20% Gránulo
Peso	1. 250 g
	2. 500 g
	3. 750 g
Velocidad	1. 60
	2. 70
	3. 80
Ciclos	1. 500
	2. 750
	3. 1000

La relación entre variables es un aspecto importante antes de realizar el entrenamiento. Esta relación identifica las mejores variables predictores para la red neuronal, y también determina la colinealidad entre variables. Los coeficientes de correlación determinan la relación entre variables mediante un valor numérico. Esta correlación se presenta con un valor entre -1 y 1 , de manera que -1 representa una correlación negativa perfecta, a medida que crece la variable A, la variable B tiende a encogerse, y 1 determina una correlación perfecta. La ausencia completa de relación se presenta cuando el coeficiente de correlación es cero. La Figura 22 presenta la correlación entre las variables independientes con respecto a la variable objetivo.

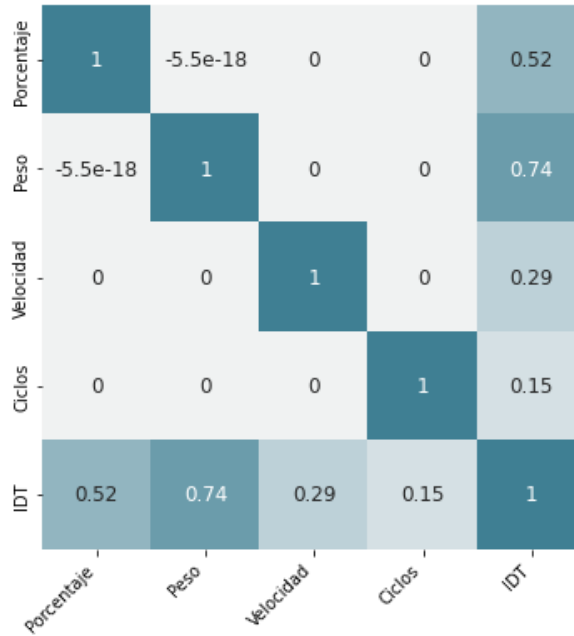


Figura 22: El coeficiente de correlación que mejor valor presenta fue entre la relación de IDT y el peso aplicado con una correlación de 0,74, mientras que la correlación más baja corresponde al IDT y los ciclos de desgaste con un valor de 0,15.

La Figura 23 representa la distribución asimétrica positiva del comportamiento de la variable de salida IDT. Presenta una cola positiva, debido a que varios datos presentan valores muy superiores a la media.

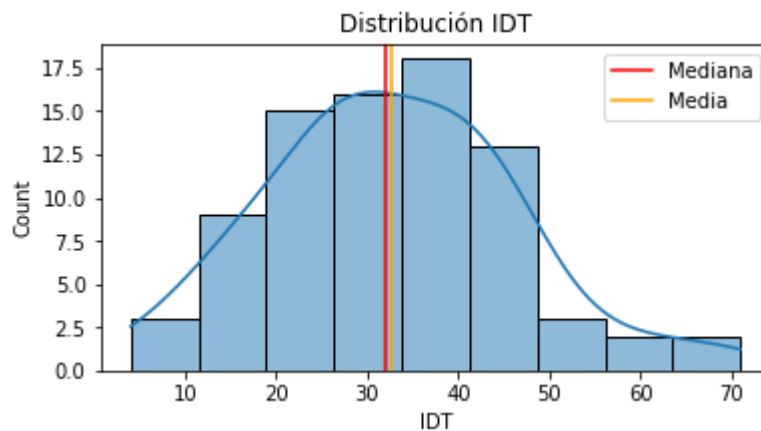


Figura 23: Distribución muestral de la variable de salida IDT caracterizada por una forma curva asimétrica positiva respecto a un valor central que coincide con la media y mediana de la distribución, la cual se extiende en la dirección positiva.

La subdivisión del conjunto de datos fue del 80% como conjunto de entrenamiento y 20% como conjunto de validación. Los datos de entrada se normalizaron antes del entrenamiento para mejorar el aprendizaje automático. La normalización de los datos de entrada es recomendable antes del proceso de entrenamiento para obtener buenos resultados, así como para agilizar significativamente los cálculos. El modelo converge sin la necesidad de normalización de datos, pero esto dificulta el entrenamiento y provoca que el modelo final dependa de la elección adecuada de unidades en la entrada. Este análisis se realizó con dos capas ocultas, cada una compuesta por 192 neuronas. Los demás hiperparámetros son: tasa de aprendizaje de 0,001, el optimizador fue RMSprop, MAE como función de pérdida, ReLu como función de activación y 500 épocas.

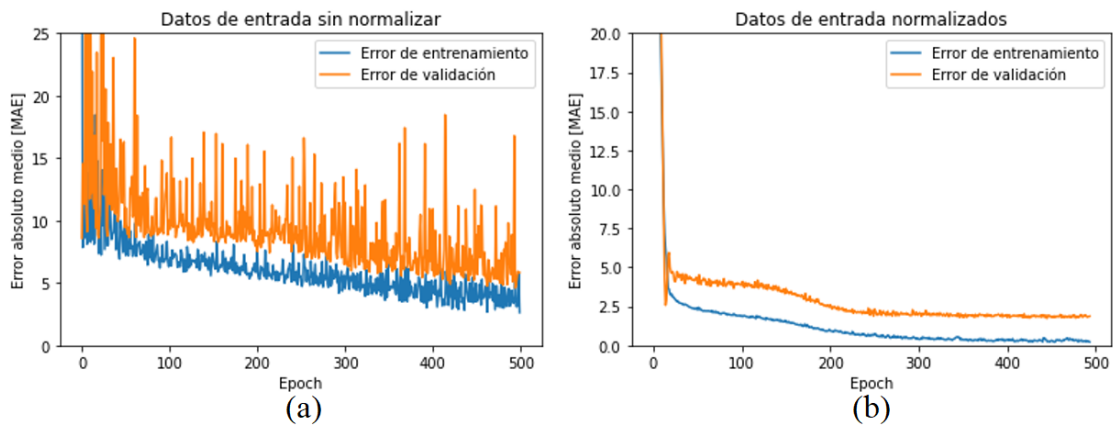


Figura 24: En el apartado (a) se muestra el error de entrenamiento y validación con los datos sin normalizar, el cual presenta muchas oscilaciones. En el apartado (b) se utilizó los datos normalizados, mejorando su estabilidad y desempeño.

La Figura 24 muestra el error de entrenamiento y validación donde se señala la diferencia entre el entrenamiento con y sin normalización. Muchos modelos neuronales son sensibles a la normalización. Esto es debido a que los datos que se midan en una escala mayor o que presenten mayor varianza dominarán el modelo, aunque no presenten la mayor relación con la variable de salida. La normalización es una forma de transformar los datos de manera que todas las variables estén aproximadamente en la misma escala. La normalización se establece como el cociente entre la diferencia del dato de entrada con su media y la desviación estándar del conjunto de datos

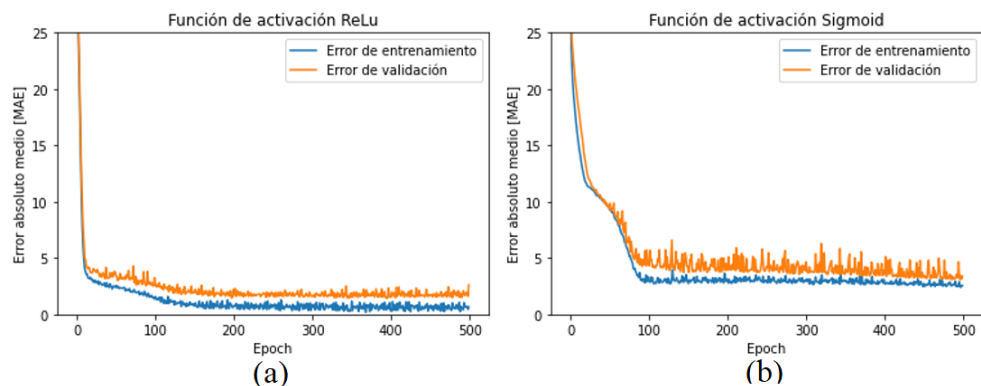
$$z = \frac{x - \mu}{\sigma}, \quad (40)$$

donde x es un dato, μ es la media y σ es la desviación estándar. En este proyecto, solo se normalizaron los datos de entrada. El IDT no fue normalizado por ser el valor esperado por el entrenamiento. La selección adecuada de hiperparámetros provoca una convergencia rápida durante el entrenamiento y una buena precisión para generar predicciones. La principal característica de este proceso es que no existe una regla general para determinar sus valores. El enfoque de prueba y error es el más utilizado para los hiperparámetros.

4.4.1. Función de activación y optimizador

Se realizó un análisis cruzado sobre el optimizador y la función de activación para determinar su relación. La arquitectura del modelo utilizado fue con una red multicapa. Se consideraron los siguientes valores de hiperparámetros:

- Dos capas ocultas (cada una con 192 neuronas),
- Número de época = 500.
- Tasa de aprendizaje = 0,001,
- Función de pérdida = MAE.



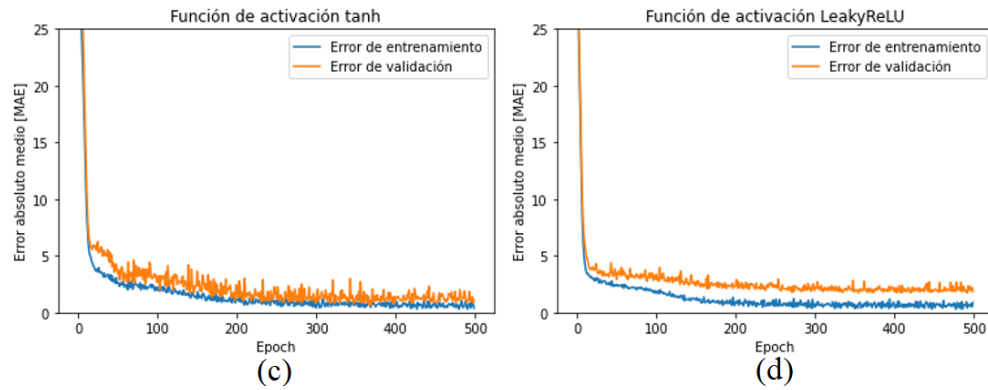


Figura 25: Error de entrenamiento y validación con el algoritmo de optimizado RMSprop. (a) ReLu desempeño buenos resultados con bajas oscilaciones. (b) Sigmoid presentó muchas oscilaciones durante todo el entrenamiento sin llegar a converger. (c) Tanh demostró oscilaciones muy bruscas durante todo el proceso. (d) LeakyReLU obtuvo buenos resultados, pero con oscilaciones pequeñas.

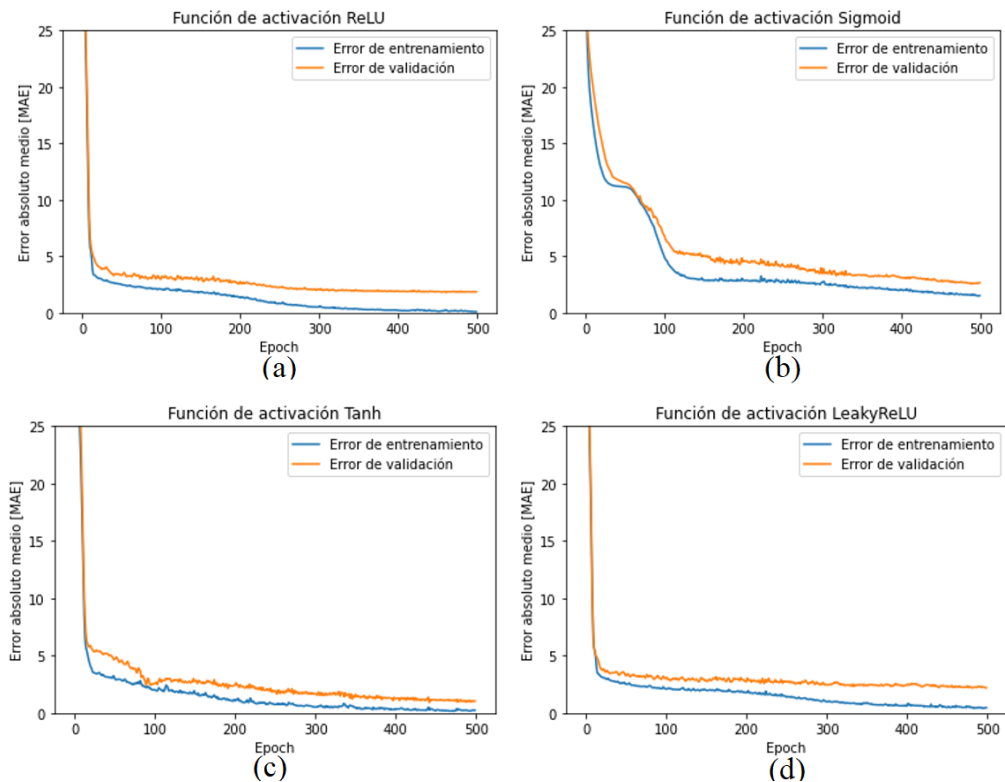


Figura 26: Error de entrenamiento y validación con el algoritmo de optimizado Adam. (a) ReLu demostró un buen desempeño y bajas oscilaciones. (b) Sigmoid empieza a converger a partir de la época 100. (c) Tanh presentó bajas oscilaciones, pero muy inestables. (d) LeakyReLU obtuvo oscilaciones más grandes con respecto a ReLU.

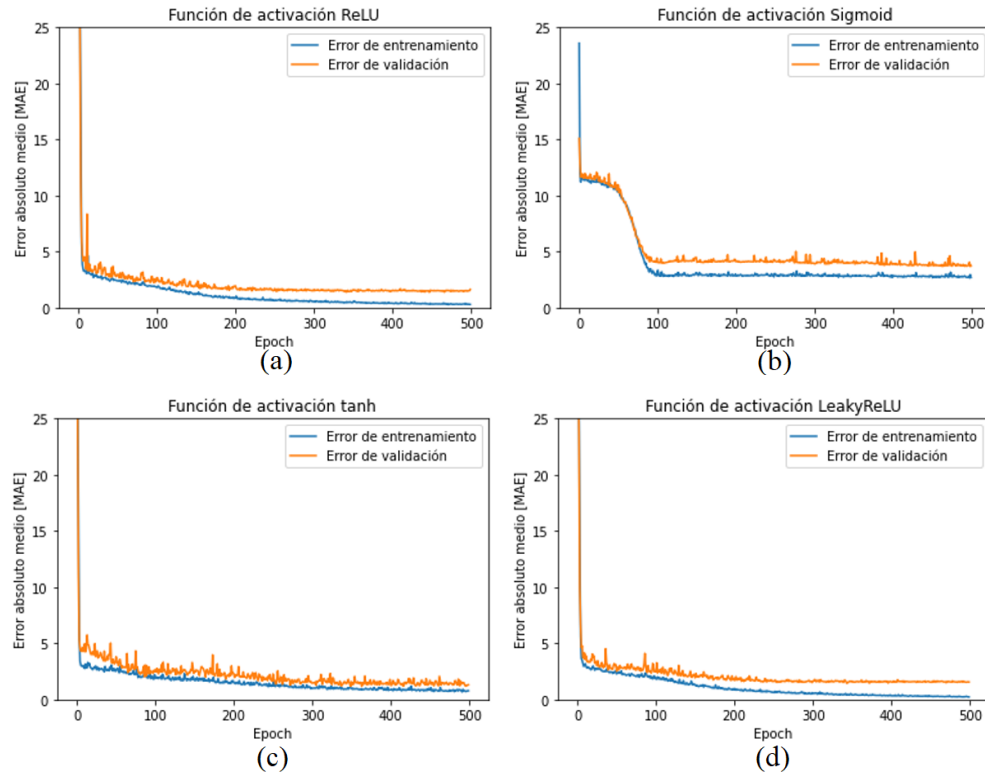


Figura 27: Error de entrenamiento y validación con SGD. (a) ReLU presenta buenos resultados con pocas oscilaciones al inicio del entrenamiento. (b) Sigmoid converge hasta la época 100 con valores elevados. (c) tanh muestra grandes oscilaciones durante el entrenamiento. (d) LeakyReLU obtuvo resultados similares a ReLU.

El objetivo de este análisis fue comprender el comportamiento del modelo al combinar varios optimizadores y funciones de activación. En base a la teoría, la mejor combinación podría ser la función de activación ReLU y el optimizador Adam. La combinación de SGD y la alta tasa de aprendizaje inicial conduce a oscilaciones visibles. El optimizador Adam tiene las oscilaciones más bajas en todos los entrenamientos. La función de activación sigmoidea requirió más épocas para que se llegara a la convergencia. ReLU y Leaky ReLU funcionan mejor con respecto a las funciones de activación sigmoidea y tanh. Se caracterizan por una convergencia rápida en cuanto al número de épocas. Por esta razón, los análisis posteriores se los realizará con el optimizador Adam, ya que fue el optimizador que presentó los mejores resultados sin cambios aparentes. El optimizador Adam tiene beneficios con respecto a SGD y RMSprop. Los beneficios de Adam ayudan a superar el problema del mínimo local.

4.4.2. Tasa de aprendizaje y Batch size

El siguiente análisis está relacionado con el Batch size y la tasa de aprendizaje del modelo neuronal. El objetivo de este análisis es extraer la relación entre los casos extremos de baja tasa de aprendizaje / alto Batch size y alta tasa de aprendizaje / bajo Batch size. Para el Batch size y la tasa de aprendizaje, se consideran los siguientes valores:

- Tres valores de Batch size (BS) diferentes (BS = 1, BS = 13, BS = 32),
- Tres valores diferentes de tasa de aprendizaje (LR) (LR = 0,001, LR = 1, LR = 10).
- Dos capas ocultas (cada una con 192 neuronas),
- Optimizador = Adam,
- Función de activación = ReLU.

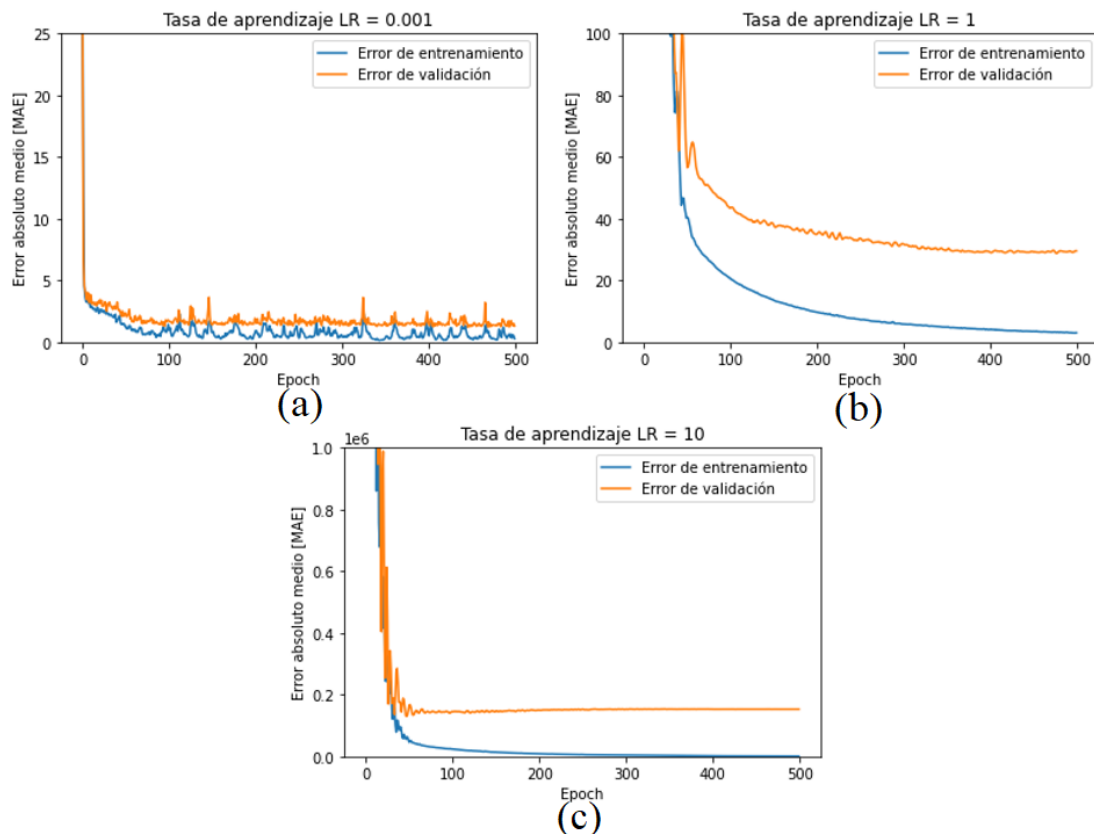


Figura 28: Error de entrenamiento y validación con BS = 1. (a) Presentó oscilaciones durante todo el entrenamiento. (b) y (c) mostraron valores muy elevados para el error de entrenamiento, dejándolos fuera del análisis.

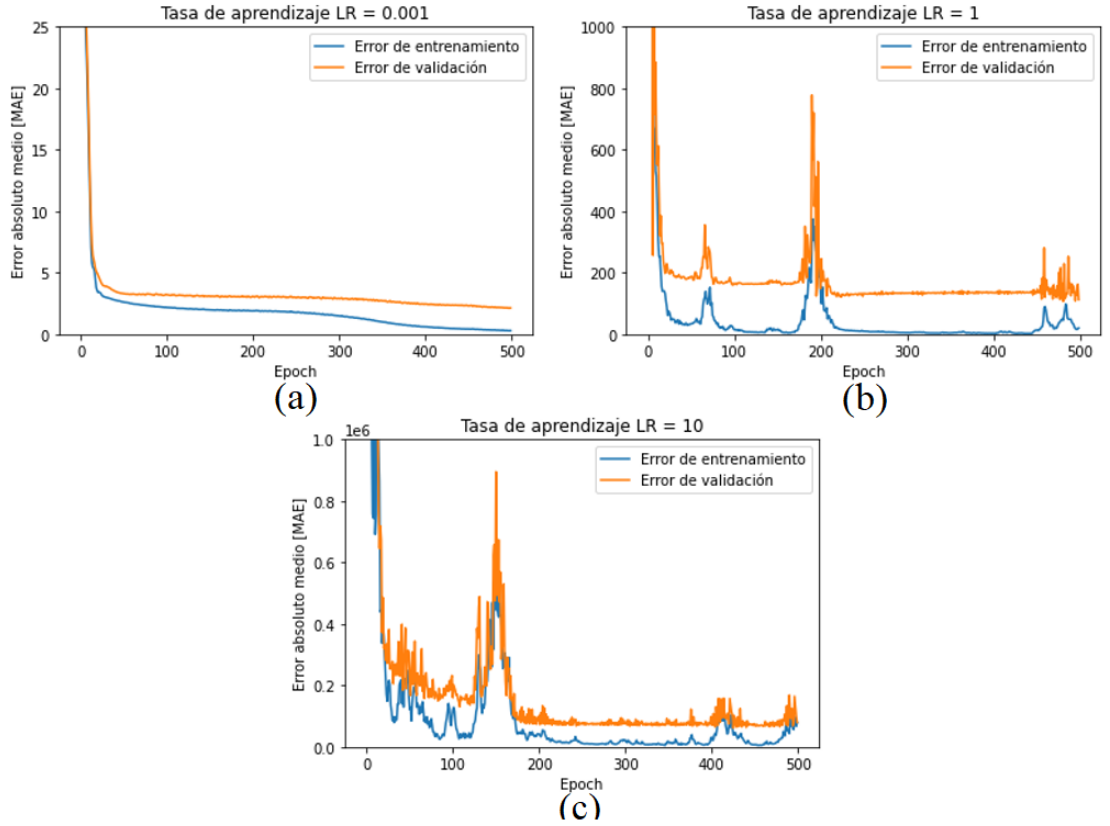
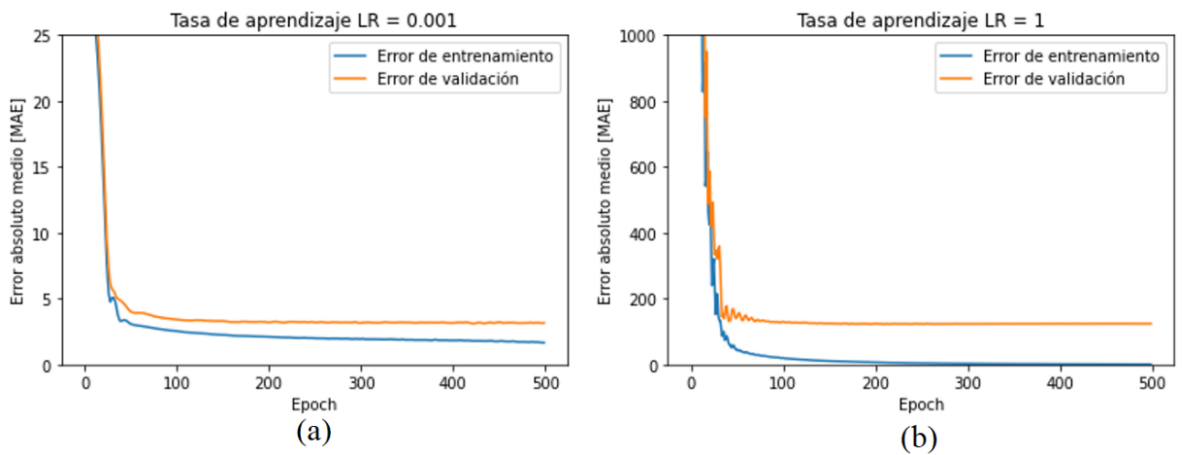


Figura 29: Error de entrenamiento y validación con BS = 13. (a) Presentó muy pocas oscilaciones y un buen desempeño. (b) y (c) mostraron valores elevados para el error de entrenamiento.



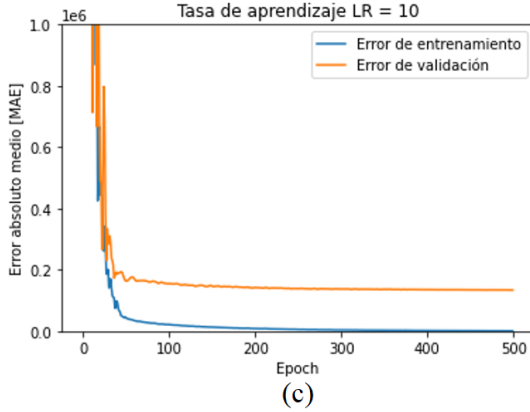


Figura 30: Error de entrenamiento y validación con $BS = 32$. (a) Presentó pocas oscilaciones y un buen desempeño. (b) y (c) mostraron valores altos para el error de entrenamiento.

Un Batch size pequeño ofrece un error de generalización menor. Los parámetros del modelo se actualizan luego de cada muestra de datos. La red pretende determinar la coherencia entre cada muestra, lo que conduce a una generalización menor. Se observa que con el LR de 0,001 las curvas de entrenamiento y validación parten de un valor menor con respecto a LR de 1 y LR de 10. Los números de actualizaciones de parámetros en una época ($N_{updates}$) se representan como el número de muestras utilizadas durante el entrenamiento dividido por el Batch size:

$$N_{updates} = \frac{N_{muestras}}{BS}. \quad (41)$$

El BS de 13 con LR de 0,001 presentaron los mejores resultados durante el proceso de aprendizaje. Esta combinación presentó el menor error de entrenamiento y validación. El LR de 10 no permitió acercarse a los mínimos globales. Los problemas de oscilación pueden considerarse resueltos cuando se utilizan Batch size de menor tamaño, lo que provoca una mejor capacidad de generalización. No existe una relación exacta entre el Batch size y la cantidad de muestras involucradas en el proceso de aprendizaje. Se establece un valor intermedio de BS de 13 en los siguientes análisis. En base a la Figura 29 es posible concluir que con un BS de 13 y LR de 0,001 (o menor) son opciones adecuadas para lograr un buen nivel de precisión.

4.4.3. Épocas y sobreajuste

El número de épocas se define como el número de veces que el error de entrenamiento se alimenta de la red. La selección del número de épocas adecuado puede conllevar a dos posibles errores frecuentes de entrenamiento: el overfitting y underfitting. La Figura 31 muestra un ejemplo drástico de sobreajuste en el entrenamiento de la red neuronal.

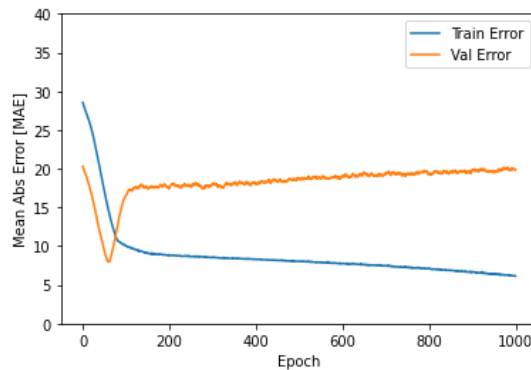


Figura 31: Modelo con overfitting, donde el error de entrenamiento disminuye mientras el error de validación aumenta.

El overfitting surge cuando el modelo continúa aprendiendo de los datos de entrenamiento. El modelo ajusta demasiado los ejemplos de entrada y no permite realizar las predicciones correctas al presentarle nuevos datos. En la Figura 31 el modelo empieza a sobreentrenar los datos alrededor de las 100 épocas. Luego de ese punto, el error de validación es visiblemente más alto que el error de entrenamiento.

La detención anticipada (EarlyStopping) es un método que permite especificar una cantidad arbitraria de épocas de entrenamiento y detener el entrenamiento una vez que el rendimiento del modelo deja de mejorar. La Figura 32 muestra el error de entrenamiento con y sin EarlyStopping. Para este análisis se determinó los siguientes hiperparámetros: función de activación ReLu, tasa de aprendizaje de 0,001 y número de épocas de 1000. Los datos de entrenamiento se tomaron en base a la validación cruzada de k-fold.

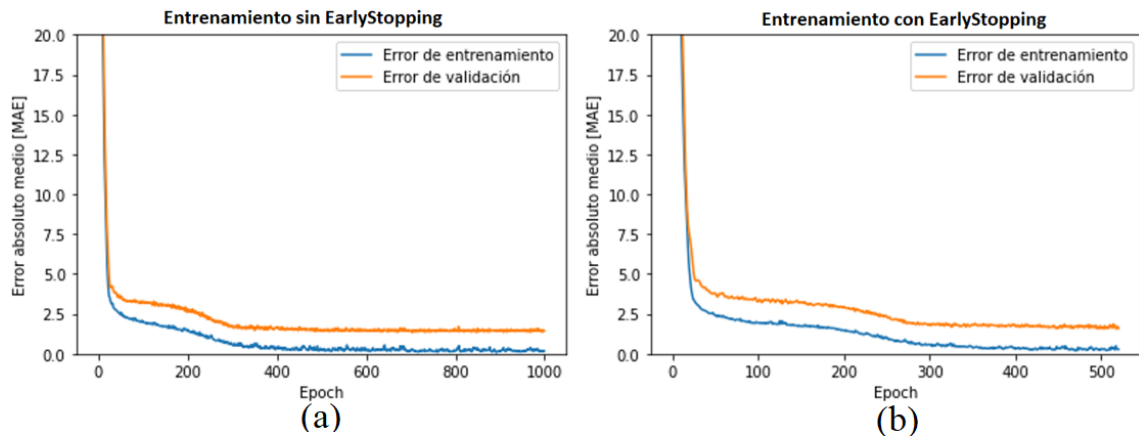


Figura 32: Análisis de overfitting del modelo neuronal. En el apartado (a) se realizó el entrenamiento sin utilizar EarlyStopping, demostrando que el número de épocas llegó a su máximo de 1000, mientras que el apartado (b) utilizando EarlyStopping el entrenamiento se detuvo en 520 épocas aproximadamente.

Se consideró una gran cantidad de épocas con el fin de señalar el comportamiento de aprendizaje hasta que la tendencia se mantenga constante. Tomando en cuenta la Figura 32 se determinó que el número de épocas adecuada sin que se produjera un overfitting ni underfitting fue de 500 épocas.

El underfitting ocurre cuando el modelo no aprende de los datos o existen muy pocas neuronas en las capas ocultas para detectar correctamente las señales en un conjunto de datos complicado. El análisis para detectar el underfitting se llevó a cabo tomando el 20%, 50% y 80% de los datos de entrenamiento. La estructura del modelo para este análisis incluyó:

- Dos capas ocultas de 192 neuronas,
- Optimizador de Adam,
- Función de activación de ReLU,
- LR = 0,001,
- BS = 13,
- Número de épocas = 500

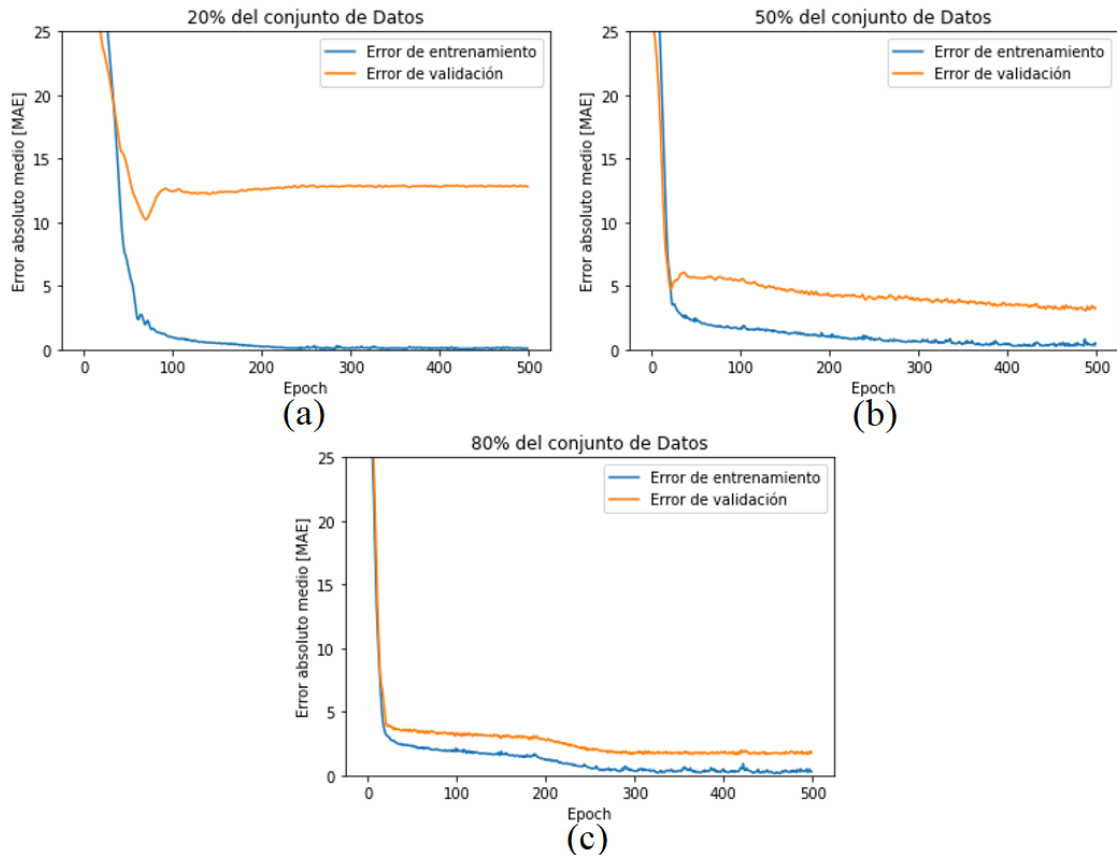


Figura 33: Análisis de underfitting del modelo neuronal. (a) Al utilizar el 20% de los datos el error de validación supera el error de entrenamiento. (b) Al entrenar el modelo con el 50% disminuye el error, pero no presenta una buena predicción. (b) Presenta el mejor error de entrenamiento con el 80% de los datos.

El aumento del tamaño del conjunto de datos de entrenamiento mejora los resultados de predicción. Las muestras utilizadas para el entrenamiento y la validación se eligieron al azar. El aumento de estos datos reduce la brecha entre el entrenamiento y el error de validación. Una vez que la curva de validación ha alcanzado su mínimo esta permanece plana.

La Tabla 18 presenta una serie de pruebas utilizando diferentes arquitecturas neuronales para determinar la mejor combinación de número de capas ocultas y neuronas para el entrenamiento de la red neuronal. Este análisis se determinó con los siguientes hiperparámetros: función de activación ReLu, batch size de 13, tasa de aprendizaje de 0,001 y número de épocas de 500.

Tabla 18: Validación estadística conforme a diferentes configuraciones neuronales.

Arquitectura	Conjunto de Datos	MAE	MSE	R ²
6-64-1	Entrenamiento	2,407	9,716	0,818
	Validación	3,253	17,364	
6-192-1	Entrenamiento	1,665	2,796	0,887
	Validación	5,285	10,762	
6-256-1	Entrenamiento	1,097	2,321	0,927
	Validación	2,538	6,932	
6-64-64-1	Entrenamiento	1,444	3,205	0,838
	Validación	4,546	15,491	
6-192-192-1	Entrenamiento	0,749	1,607	0,961
	Validación	1,198	3,690	
6-256-256-1	Entrenamiento	0,513	1,740	0,959
	Validación	0,788	3,885	
6-64-64-64-1	Entrenamiento	0,897	1,520	0,956
	Validación	2,140	3,227	
6-192-192-192-1	Entrenamiento	0,888	2,187	0,930
	Validación	1,357	6,667	
6-256-256-256-1	Entrenamiento	1,134	2,044	0,944
	Validación	2,312	5,324	

Las mejores redes con topologías preseleccionadas son aquellas con dos capas ocultas. La arquitectura con mayor eficiencia en su pronóstico fue 6-192-192-1, es decir, 6 neuronas de entrada, 192 neuronas en 2 capas ocultas y una neurona en la capa de salida. Esta red presentó un R² de 0,961, es decir, es capaz de predecir la variable de salida con una eficiencia del 96,1%.

La validación cruzada es una práctica común que se realiza en el aprendizaje supervisado para evitar un ajuste excesivo. Cuando un algoritmo está configurado para entrenar en aprendizaje supervisado, una parte del conjunto de datos se mantiene para probar el rendimiento final del modelo. Antes del proceso de aprendizaje, se realiza una subdivisión de todo el conjunto de datos. Esto reduce drásticamente el número de muestras utilizadas en el entrenamiento y conduce a una alta dependencia de los datos aleatorios considerados. El enfoque básico de la validación cruzada es la subdivisión del conjunto de entrenamiento en k veces. Un conjunto se mantiene para la evaluación final (conjunto de prueba) y el resto $k-1$ conjuntos se utilizan como conjunto de entrenamiento. Esto reduce el sesgo en el entrenamiento y minimiza la varianza en la validación. Cada pliegue se entrena una vez

y se determina el error de prueba. Los enfoques más comunes son $k = 5$ o $k = 10$. En esta investigación se consideró $k = 5$. Además, el procedimiento se repite dos veces para aumentar la validez general del resultado final.

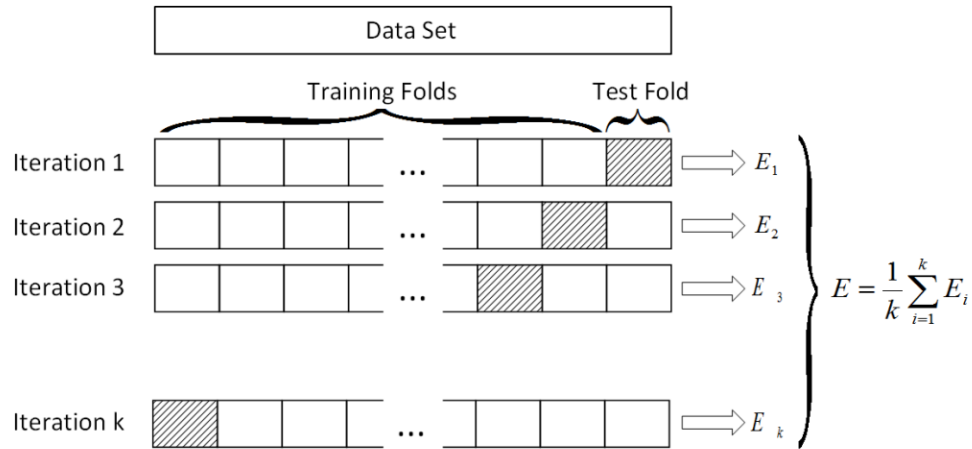


Figura 34: Validación cruzada de K-fold para evaluar el rendimiento del modelo mediante el entrenamiento de k subconjuntos de datos y evaluados en el subconjunto complementario. Fuente: [44]

4.5. Análisis de resultados

El modelo neuronal se desarrolló para estimar el índice de desgaste Taber (IDT) en base a diferentes combinaciones. La base de datos se obtuvo a través de ensayos de laboratorio. Las placas fueron diseñadas en base a la norma ISO 9352. Las variables independientes fueron: el porcentaje volumétrico, la carga aplicada, la velocidad de giro y el número de ciclos. El conjunto de datos experimentales utilizados para la etapa de entrenamiento se dividió mediante validación cruzada en: 80% para entrenamiento y 20% para validación. Se desarrolló con un esquema neuronal de feed-forward y backpropagation sobre un total de 81 datos experimentales. El modelo que presentó los mejores resultados fue con la arquitectura 6-192-192-1. Se utilizó la función de activación ReLu tanto para las capas ocultas como para la capa de salida. La Figura 35 presenta la relación entre los valores obtenidos a través del entrenamiento neuronal y los valores obtenidos de manera experimental.

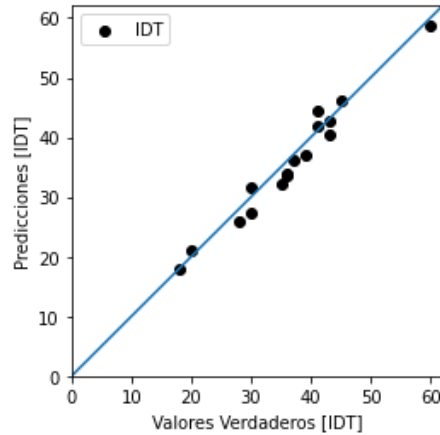


Figura 35: Regresión entre los datos presentados por la red neuronal en comparación con los valores reales obtenidos de manera experimental.

El rendimiento de predicción del modelo se evaluó mediante los valores de R^2 , MAE y MSE. En base a las investigaciones realizadas por otros autores, se recomienda que el modelo presente un valor de R^2 alto y las mediciones estadísticas MAE y MSE sean bajas para que el modelo neuronal tenga una alta eficiencia. La Tabla 19 presenta los resultados del MAE y MSE luego del entrenamiento de la red. Los resultados estadísticos de MSE y MAE están dentro de un rango aceptable y cumplen con la integridad de las etapas de aprendizaje y prueba de la red.

Tabla 19: Resultados estadísticos de MAE, MSE y R^2 de modelo neuronal.

Conjunto de datos	MAE	MSE	R^2
Entrenamiento	0,749	1,607	0,961
Validación	1,198	3,690	

La Tabla 20 presenta la mejor configuración de hiperparámetros para el modelo neuronal, tomando en cuenta los resultados estadísticos y el costo computacional.

Tabla 20: Hiperparámetros del modelo neuronal.

Parámetro	Valor
Arquitectura de la Red	6-192-192-1
Parámetros de entrada	<ul style="list-style-type: none"> • Porcentaje volumétrico • Carga aplicada • Velocidad de giro • Ciclos
Parámetros de salida	Índice de desgaste Taber (IDT)
Función de activación	ReLu
Optimizador	Adam
Función de rendimiento	<ul style="list-style-type: none"> • Error cuadrático medio (MSE) • Error absoluto medio (MAE) • Coeficiente de determinación (R^2)
Tasa de aprendizaje	0,001
Número de iteraciones	500
Batch size	13

La Figura 36 representa la gráfica de rendimiento de la red a través de los errores de entrenamiento y validación. Este análisis se realizó tomando en cuenta los hiperparámetros establecidos en la Tabla 20 y se analizaron el MAE y MSE.

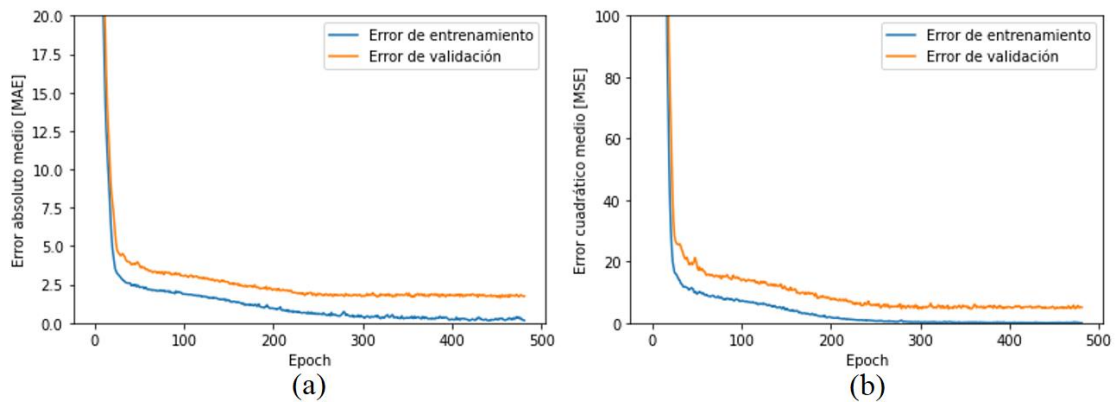


Figura 36: Error de entrenamiento y validación de la red. (a) Representa el MAE obtenido por la red, donde se obtuvo un valor de 0,749 mientras que el error de validación fue de 1,198. (b) Representa el MSE obteniendo por la red, presentando un valor de 1,607 y el error de validación se obtuvo 3,69.

La Figura 37 representa la distribución de residuos del modelo. La mayor cantidad de residuos se acercan al valor de cero, por lo que también determina el buen desempeño del modelo neuronal propuesto. Los residuos no se aproximan a una curva normal. Este comportamiento se debe al pequeño tamaño muestral considerado para el entrenamiento.

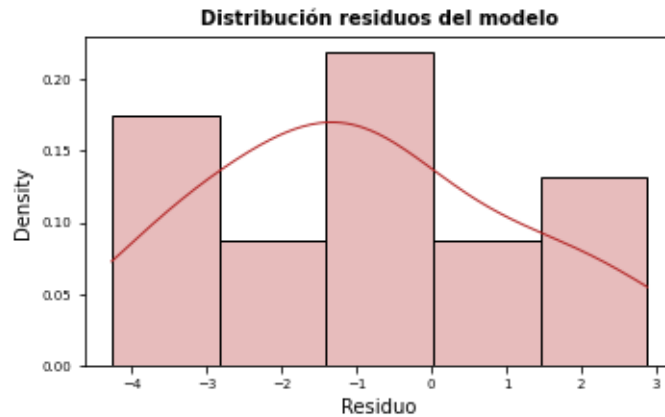


Figura 37: La grafica de distribución de residuos tiende a una curva gaussiana, es decir, la mayoría de los residuos tienden a acercarse al valor de cero.

La predicción de la red se determinó a través de los valores estimados durante la validación del modelo neuronal. Estos valores no se utilizaron durante el proceso de entrenamiento. La figura 38 representa la predicción de la red neuronal en comparación de los valores reales. Sin embargo, se observa que varios valores no concuerdan con los datos reales. Este problema pudo ser provocado por los siguientes factores: tipo de material, falta de precisión en la toma de muestras y la arquitectura de la red neuronal.

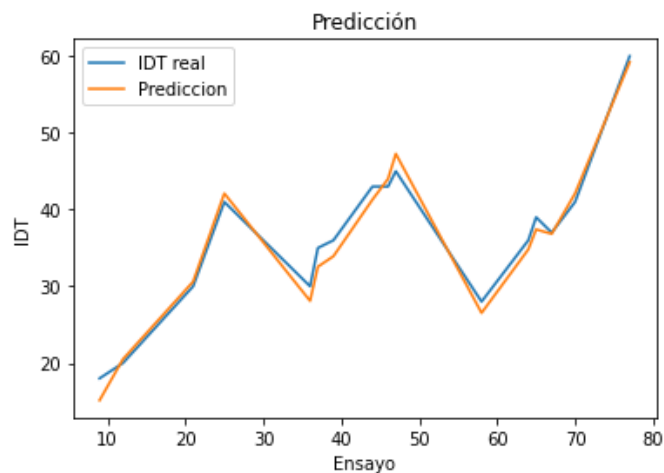


Figura 38: Datos predichos por el modelo neuronal toman la forma de los datos obtenidos de manera experimental.

La Figura 39 representa el porcentaje de error presentado por la red en cada dato de validación. En base a esta figura se considera una correlación aceptable, debido a que el error promedio entre los valores de respuesta reales y los predichos por el modelo neuronal

fue de 4,418 %. Los resultados indican las similitudes entre el estudio experimental y el modelo neuronal, apoyando la confiabilidad del modelo.

IDT real	Prediccion	Error	%Error
18	15.145455	2.854545	15.858581
20	20.508560	-0.508560	2.542801
30	30.610209	-0.610209	2.034028
41	42.105713	-1.105713	2.696861
30	28.101835	1.898165	6.327216
35	32.548637	2.451363	7.003893
36	33.922871	2.077129	5.769804
43	41.292782	1.707218	3.970275
43	44.002468	-1.002468	2.331321
45	47.271732	-2.271732	5.048294
28	26.527536	1.472464	5.258799
36	34.812180	1.187820	3.299501
39	37.409454	1.590546	4.078322
37	36.838505	0.161495	0.436474
41	42.114059	-1.114059	2.717218
60	59.213940	0.786060	1.310101

Figura 39: Porcentaje del error de predicción en comparación con los valores reales.

El análisis de sensibilidad representa el efecto que provoca la variación de un parámetro de entrada sobre la variable de salida. La Figura 40 muestra la correlación de Pearson de las variables de estudio. El índice de análisis de sensibilidad (SAI) está delimitado por -1 y $+1$, donde $+1$ denota el efecto proporcional máximo de una variable de entrada en el parámetro de salida, mientras que -1 demuestra el impacto inverso máximo. Un valor de cero representa la ausencia de correlación entre la salida y la entrada.

index	variable_1	variable_2	r	abs_r
21	IDT	Peso	0.7354229158270836	0.7354229158270836
9	Peso	IDT	0.7354229158270836	0.7354229158270836
4	Porcentaje	IDT	0.5174610866288473	0.5174610866288473
20	IDT	Porcentaje	0.5174610866288473	0.5174610866288473
14	Velocidad	IDT	0.2863726953779303	0.2863726953779303
22	IDT	Velocidad	0.2863726953779303	0.2863726953779303
23	IDT	Ciclos	0.154796051555638	0.154796051555638
19	Ciclos	IDT	0.154796051555638	0.154796051555638
1	Porcentaje	Peso	-5.513000416713933e-18	5.513000416713933e-18
5	Peso	Porcentaje	-5.513000416713933e-18	5.513000416713933e-18

Figura 40: La correlación de Pearson entre variables presenta una relación positiva entre el IDT y el peso, mientras que la peor la relación corresponde al porcentaje volumétrico y el peso aplicado.

Los resultados demuestran una fuerte correlación positiva entre la variable de salida y el peso aplicado sobre la placa. Este resultado determina que a medida que aumenta el peso aplicado sobre la placa, también aumenta el porcentaje volumétrico. La correlación entre el porcentaje volumétrico y el peso aplicado es cercana a cero, lo que implica una baja relación entre variables. Esta correlación se la considera insignificante. La Figura 41 representa los residuos obtenidos por la red neuronal en relación con los valores predicción de manera experimental. Los residuos se distribuyen aleatoriamente, determinando un supuesto de homocedasticidad u homogeneidad de varianzas. Se afirma que el modelo generaliza de manera aceptable la variable de salida.

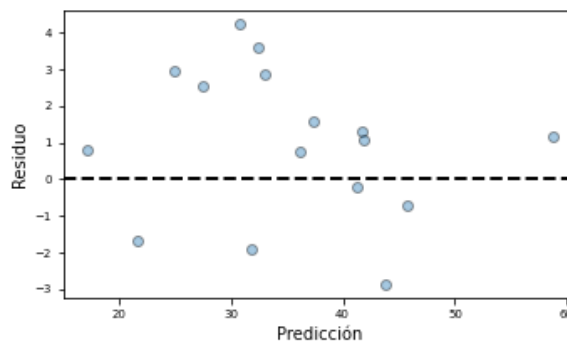


Figura 41: La dispersión de residuos distribuidos de forma homogénea. El patrón de residuos se distribuye al azar, no presenta tendencias ni tampoco dispersiones.

La normalidad del modelo se realizó a través de la prueba Q-Q. La Figura 42 hace referencia a los cuantiles del modelo. En esta figura se muestra que la mayoría de los errores residuales se encuentran sobre la línea de tendencia, es decir, tienden a seguir una distribución normal.

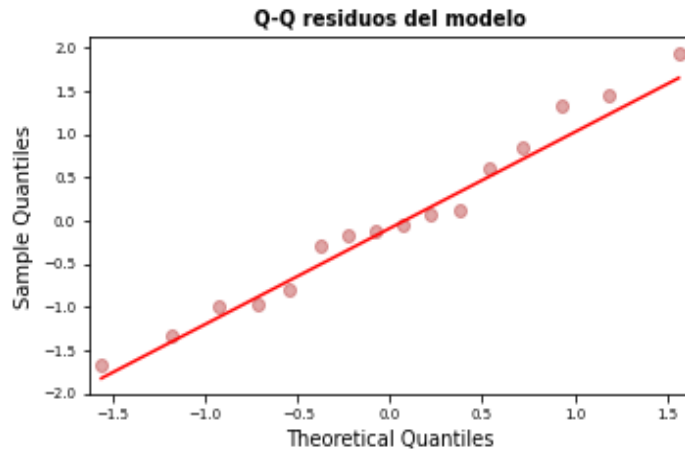


Figura 42: Gráfico de cuantiles teóricos del modelo. Los residuos presentan una distribución normal a lo largo de la línea de tendencia.

4.6. Discusión

En esta investigación se estudió y analizó el alcance de la integración de técnicas de Machine Learning (ML) a la predicción del desgaste abrasivo en placas de látex y partículas de caucho reciclado, en base a diferentes configuraciones volumétricas. Para cumplir con el objetivo del proyecto, se analizaron diferentes modelos basados en redes neuronales entrenados de manera supervisada. Se utilizó un conjunto de 81 datos, generados de manera experimental. La aplicación de algoritmos de aprendizaje automático en los sistemas de predicción de desgaste abrasivo ha sido estudiada por varios años. Es una aplicación que aún no se consolida y presenta grandes aportaciones a la industria. A partir de los resultados obtenidos, se rechaza la hipótesis nula y se acepta la hipótesis alternativa, la cual establece que con los parámetros de entrada es posible determinar el desgaste abrasivo en un porcentaje mayor del 90%.

Estos resultados guardan relación con varias investigaciones, tal como estableció B. Shuvho y M. Chowdhury, en su investigación [45], donde utilizaron RNA para predecir la tasa de erosión del polímero de nailon-12 y garantizar la precisión de la computación blanda mediante un algoritmo de backpropagation. Las entradas incluyeron la velocidad de impacto, ángulo de impacto, tamaño de erosión y distancia de separación. La consistencia entre los valores del modelo experimental y el modelo neuronal se estableció con un coeficiente de determinación R^2 del 94,428% y el error cuadrático medio de la raíz de 9,729. En la presente investigación se obtuvo un coeficiente de determinación del 97,3% y un error cuadrático medio de la raíz de 1,920. Lo que significa que el modelo de red propuesto tiene un alto grado de confiabilidad para predecir el desgaste abrasivo. Z. Zhang y M. Barkoula, [28], aplicaron una RNA para predecir el desgaste erosivo de tres polímeros: polietileno (PE), poliuretano (PUR) y epoxi modificado por poliuretano descompuesto higrotérmicamente (EP-PUR). Se utilizaron tres conjuntos de datos independientes de mediciones de desgaste erosivo y propiedades características de estos polímeros para entrenar y probar las redes neuronales. Se alcanzaron resultados de predicción aceptables, y obtuvo un coeficiente de determinación $R^2 \geq 0,9$ para los tres casos. En base a los resultados obtenidos por esta investigación se ratifica la validez del modelo neuronal.

K. Veltena y R. Reinicke en su estudio, [26], determinaron el potencial de las técnicas de machine learning para predecir y analizar el comportamiento de desgaste de materiales de cojinetes poliméricos reforzados con fibra corta. Se utilizaron un mayor número de variables de entrada que caracterizaban los materiales y las condiciones experimentales. Se implementó una técnica estadística para mejorar las capacidades predictivas de la red neuronal artificial (regularización bayesiana en lugar de la parada temprana). La investigación concluyó que, para obtener un coeficiente de determinación alto, es decir, por encima de 90%, se debe entrenar la red con un gran número de datos experimentales. Sin embargo, se confirmó que los resultados obtenidos son aceptables en comparación con el error sustancial de medición del volumen de desgaste. Además, se podría esperar mejores resultados en el futuro a partir de una mayor optimización del modelo neuronal,

así como una mayor disponibilidad de datos de medición para el entrenamiento de la red. Por otro lado, la investigación desarrollada por Ganesh Radhakrishnan [13] y Jairo Castillo-Calderón [37], determinaron que mientras menos neuronas tenga la capa oculta mejor rendimiento presentará el modelo, esto se puede contradecir en esta investigación, ya que según la Tabla 18 los mejores resultados de R^2 se obtuvieron con un mayor número de neuronas en la capa oculta.

Esta investigación no solo pretende construir potentes instrumentos de predicción mediante redes neuronales, sino que también analizar el impacto de las variables predictoras en cualquier conjunto de datos. El trabajo futuro de este proyecto se centrará en replicar los resultados obtenidos mediante el uso de otras técnicas computacionales que establezcan escenarios diferentes a los que se utilizaron en este estudio. Se observará el efecto de manipular el grado de relación entre las variables de entrada y salidas sobre la eficacia de los métodos propuestos.

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

- Se exploró el uso de las RNAs para estudiar el efecto del porcentaje volumétrico, la carga aplicada, la velocidad de giro y el número de ciclos sobre la pérdida de volumen en la superficie de placas de látex con partículas de caucho reciclado. El modelo neuronal generó resultados satisfactorios en comparación con las mediciones experimentales. El error absoluto medio (MAE) para los valores predichos no superó el 3,69 de promedio. Por lo tanto, mediante el uso las RNAs, se puede obtener resultados satisfactorios para predecir el desgaste, lo que a su vez permitirá estimar en lugar de medir, reduciendo el tiempo y el costo de pruebas futuras. Las placas con una composición volumétrica 60-40 presentaron la mayor resistencia al desgaste por abrasión. La resistencia al desgaste crecía proporcionalmente con el incremento del contenido de partículas de caucho reciclado en las placas.
- Se desarrolló una RNA del tipo feedforward, con los algoritmos de entrenamiento backpropagation, algoritmo de optimización Adam y con una topología final de aplicación de: 6 neuronas de entrada (3 variaciones volumétricas de la placa, la velocidad de giro, carga aplicada y el número de ciclos.), 2 capas ocultas de 256 neuronas cada una, y 1 capa de salida (IDT), con un entrenamiento de 500 etapas o ciclos, tasa de aprendizaje de 0,001, batch-size de 13, se logró predecir satisfactoriamente los valores correspondientes a la variable de salida, obteniendo un error global promedio de 4,418% y coeficientes de determinación R^2 mayores a 0,95 en la comparación de los valores experimentales con los valores predichos.
- Se investigó los algoritmos de aprendizaje automático más utilizados de las redes neuronales, en términos de aplicabilidad y restricciones, tomando en cuenta el campo de aplicación, que en este estudio fue la predicción de desgaste abrasivo. Para determinar dicha predicción se tomaron 3 condiciones de composición volumétrico, se manejaron datos normalizados y se probó diferentes configuraciones de datos de entrada. Se realizó un análisis cruzado de las

funciones de activación y los optimizadores demostrando que ReLu / Leaky ReLu con el optimizador Adam se puede considerar como la mejor combinación de función de activación y optimizador. El optimizador Adam permite una convergencia rápida y bajas oscilaciones. Los valores bajos de batch size se definieron como ruidosos y se caracterizaron por la presencia de oscilaciones. Adicionalmente, cuando se consideraron valores altos de tasa de aprendizaje, las oscilaciones eran mucho más altas y el error de validación presentaba oscilaciones bruscas. El número de épocas se ajustó correctamente cuando se alcanzó el error mínimo de validación y no provocó un sobreajuste del modelo. El modelo se ajustó con una precisión alta del 95%, pero esto no conduce a una menor capacidad de generalización. La validación cruzada fue una herramienta útil para analizar la capacidad de generalización. Los modelos con mejor capacidad de generalización fueron los que tienen dos o tres capas ocultas. El conocimiento general adquirido de base a artículos anteriores demuestran que las estructuras complejas funcionan mejor en este tipo de estudios.

5.2. Recomendaciones

- En base a la presente investigación se demostró un alto grado de confiabilidad en la predicción del desgaste abrasivo del material compuesto mediante redes neuronales entrenadas con una baja cantidad de muestras. Este método podría ser utilizado por diferentes empresas que utilicen el caucho como materia prima, ya que se podrán reemplazar estos materiales convencionales o a su vez reintroducir los desechos de caucho al proceso de fabricación combinándolos con una matriz acorde a las necesidades de la empresa. De esta manera favorecerá la calidad final del producto y mejorará la rentabilidad de la empresa.
- Se recomienda implementar arquitecturas de mayor complejidad, algoritmos y enfoques diferentes, que no fueron abordados en esta tesis. En cuanto a la comprensión de redes neuronales de tipo regresión, los trabajos futuros deberán implicar una aplicación de estos métodos a problemas de mayor complejidad, que involucren problemas reales de diversas áreas de aplicación.

- En lo referente al desarrollo de las probetas es recomendable controlar el tiempo de vulcanizado, también prevenir ingreso líquido en la mezcla, debido a que provocaría el deterioro del material.

5.3. Bibliografía

- [1] G. Peláez Arroyave, «Aplicaciones de caucho reciclado: Una revisión de la literatura,» 2017.
- [2] M. Posso Cárdenas y M. Buenaño Buenaño, «Estudio del sector productivo automotriz del reciclaje de neumáticos usados para la implementación en la matriz productiva del Ecuador,» Quito, 2014.
- [3] B. Cerpa Gonzales, «Propuesta de un proceso de reciclado para optimizar el uso de residuos de caucho negro en la línea de producción de suelas en una empresa del sector calzado, Arequipa.,» Arequipa, 2019.
- [4] Y. Fang y M. Zhan, «The status of recycling of waste rubber,» *Science Direct*, 2001.
- [5] D. Soria Mejía, «Implementación de un sistema predictivo con redes neuronales,» Ambato, 2021.
- [6] A. S. Acevedo Eder y E. Serna, «Principios y características de las redes neuronales artificiales.,» *Desarrollo e innovación en Ingeniería*, p. 173, 2017.
- [7] L. Quiñones, L. Ochoa y O. Gamarra, «Red neuronal artificial para estimar un índice de calidad de agua,» *Enfoque UTE*, vol. 11, n° 2, pp. 109--120, 2020.
- [8] A. Criollo Salamea, «Caracterización de caucho reciclado proveniente de Scrap y de neumáticos fuera de uso para su potencial aplicación como materia prima.,» Cuenca, 2014.
- [9] M. Álvar González, «Aplicación del caucho reciclado como solución constructiva ecológica,» Valencia, 2015.
- [10] M. Flores, «Análisis de desgastes mecánicos por tribología para reducir costos de mantenimiento del motor de tractor sobre orugas D6T-Caterpillar,» 2017.
- [11] L. Cavaleri, P. Asteris y P. Psyllaki, «Prediction of surface treatment effects on the tribological performance of tool steels using artificial neural networks,» *Applied Sciences*, vol. 9, n° 14, p. 2788, 2019.
- [12] J. Criollo Rojas, «Influencia de la exposición a la interperie en el desgaste de un material compuesto de látex y partículas de granulo de caucho reciclado.,» Ambato, 2021.
- [13] G. Radhakrishnan, «Application of artificial neural network (ANN) for predicting the wear behaviour of Al 2219-SiCp composite.,» *Applied Mechanics and Materials*, vol. 852, pp. 397--401, 2016.
- [14] O. I. Abiodun, A. Jantan y A. Omolara, «State-of-the-art in artificial neural network applications: A survey,» *Heliyon*, vol. 4, n° 11, p. e00938, 2018.
- [15] S. Sardar, S. Dey y D. Das, «Modelling of tribological responses of composites

- using integrated ANN-GA technique,» *Journal of Composite Materials*, vol. 55, n° 7, pp. 873--896, 2021.
- [16] G. A. Flores Mendoza, «Análisis tribológico para la predicción del desgaste en acero inoxidable mediante una red neuronal artificial,» Saltillo, 2016.
- [17] Z. Zhang y K. Friedrich, «Prediction on tribological properties of short fibre composites using artificial neural networks,» vol. 252, pp. 668--675, 2002.
- [18] W. P. Herrera Castillo y O. G. Toapanta Cunalata, «Principios de tribología aplicados en la ingeniería mecánica,» *3Ciencias*, vol. 57, 2019.
- [19] M. . A. Gómez Botero, «Caracterización de las propiedades tribológicas de los recubrimientos duros,» 2005.
- [20] N. Feito, A. Muñoz Sánchez, A. Díaz y J. Loya, «Analysis of the machinability of carbon fiber composite materials in function of tool wear and cutting parameters using the artificial neural network approach,» *Materials*, vol. 12, n° 17, p. 2747, 2019.
- [21] J. F. Santa Marín, «Tribología: pasado, presente y futuro,» *TecnoLógicas*, vol. 18, pp. 09-10, 2015.
- [22] A. A. Agbeleye y D. Esezobor, «Prediction of the abrasive wear behaviour of heat-treated aluminium-clay composites using an artificial neural network,» 2018.
- [23] K. Abd El-Aziz, D. Saber y A. Megahed, «Investigation and Prediction of Abrasive Wear Rate of Heat-Treated HCCIs with Different Cr/C Ratios Using Artificial Neural Networks,» *International Journal of Metalcasting*, pp. 1--15, 2020.
- [24] . I. Kopal, Labaj, Harničárová, Valíček y Hrubý, «Prediction of the tensile response of carbon black filled rubber blends by artificial neural network,» *Polymers*, vol. 10, n° 6, p. 644, 2018.
- [25] X. LiuJie, J. P. Davim y R. Cardoso, «Prediction on tribological behaviour of composite PEEK-CF30 using artificial neural networks,» *Journal of Materials Processing Technology*, pp. 374-378, 2007.
- [26] K. Velten y R. Reinicke, «Wear volume prediction with artificial neural networks,» *Tribology International*, vol. 33, n° 10, pp. 731--736, 2000.
- [27] N. Zhang , D. Lei y F. Zhao, «An improved Adagrad gradient descent optimization algorithm,» de *2018 Chinese Automation Congress (CAC)*, 2019, pp. 2359--2362.
- [28] Z. Zhang y N. Barkoula, «Artificial neural network predictions on erosive wear of polymers,» *Wear*, vol. 255, n° 1-6, pp. 708--713, 2003.
- [29] H. Alimam y M. Hinnawi, «ANN & ANFIS Models for Prediction of Abrasive Wear of 3105 Aluminium Alloy with Polyurethane Coating,» *Tribology in industry*, vol. 38, n° 2, 2016.
- [30] V. Gopan y S. Arun, «Integrated ANN-GA approach for predictive modeling and optimization of grinding parameters with surface roughness as the response,» *Materials today: proceedings*, vol. 5, n° 5, pp. 12133--12141, 2018.
- [31] M. Wiciak, A. Felusiak-Czyryca y P. Twardowski, «Tool Wear Prediction Based on Artificial Neural Network during Aluminum Matrix Composite Milling,» *Sensors*, vol. 20, n° 20, p. 5798, 2020.

- [32] R. Laghari, J. Li y A. Laghari, «carbide tool life prediction and modeling in SiCp/Al turning process via artificial neural network approach», vol. 600, 2019, p. 012022.
- [33] Hawryluk y Mrzyglod, «A system of analysis and prediction of the loss of forging tool material applying artificial neural networks.,» *Journal of Mining and Metallurgy, Section B: Metallurgy*, vol. 54, nº 3, pp. 323--337, 2018.
- [34] Suresh, A. Joshi y Manjaiah, «Experimental Investigation on Tool Wear in AISI H13 Die Steel Turning Using RSM and ANN Methods,» *Arabian Journal for Science and Engineering*, vol. 46, nº 3, pp. 2311--2325, 2021.
- [35] A. Sagbas y F. Kahraman, «Modeling and predicting abrasive wear behaviour of poly oxy methylenes using response surface methodology and neural networks,» 2016.
- [36] A. Elsheikh, S. Sharshir y M. Elaziz, «Modeling of solar energy systems using artificial neural network: A comprehensive review,» *Solar Energy*, vol. 180, pp. 622--639, 2019.
- [37] J. Castillo-Calderón, B. Solórzano-Castillo y J. Moreno-Moreno, «Diseño de una red neuronal para la predicción del coeficiente de pérdidas primarias en régimen de flujo turbulento.,» *Ingenius. Revista de Ciencia y Tecnología*, nº 20, pp. 21--27, 2018.
- [38] D. Dhande y M. Phate, «Comparative Analysis of Abrasive Wear Using Response Surface Method and Artificial Neural Network,» *Journal of The Institution of Engineers (India): Series D*, vol. 102, nº 1, pp. 27--37, 2021.
- [39] Z. Jiang, L. A. Gyurova y A. Schlarb, «Study on friction and wear behavior of polyphenylene sulfide composites reinforced by short carbon fibers and sub-micro TiO₂ particles,» *Composites Science and Technology*, pp. 734-742, 2008.
- [40] S. Espinoza Ruiz, «Diseño mecánico eléctrico de un tribómetro para ensayos de resistencia al desgaste abrasivo en pastillas de frenos tipo disco según la norma ASTM G65.,» 2017.
- [41] W. Quintero, M. d. P. Sanabria y J. Florez, «Construcción de un equipo para desgaste abrasivo según norma ASTM G--65,» *Scientia et technical*, vol. 20, nº 3, pp. 217-224, 2015.
- [42] A. O. Aguilar Aldas y A. R. Moya Moya, «Automatización de la máquina de fricción y desgaste de cilindros cruzados del Laboratorio de Tribología de la Escuela Superior Politécnica de Chimborazo.,» 2019.
- [43] J. R. Hilera González y V. J. Martínez Hernando, «Redes neuronales artificiales: fundamentos, modelos y aplicaciones,» 2000.
- [44] R. Oliveira y R. Araújo, «A System Based on Artificial Neural Networks for Automatic Classification of Hydro generator Stator Windings Partial Discharges,» Belém-PA, 2017.
- [45] B. A. Shuvho y . A. Mohammad, «Analysis of artificial neural network for predicting erosive wear of nylon-12 polymer.,» *Materials Performance and Characterization*, vol. 8, nº 1, pp. 288--300, 2019.
- [46] M. Quadrio, «A NEURAL NETWORK APPROACH FOR PREDICTING THE AERODYNAMIC PERFORMANCE OF AIRFOILS,» Milán, 2020.

- [47] Y. Goldberg, «Neural network methods for natural language processing,» *Synthesis Lectures on Human Language Technologies*, vol. 10, n° 1, pp. 1-309, 2017.
- [48] B. A. Peralta Arboleda, «Desarrollo de una red neuronal artificial para la predicción de la tendencia corrosiva e incrustante en tuberías de producción de petróleo,» 2020.
- [49] M. Rubiolo, «Desarrollo de nuevos modelos y algoritmos basados en redes neuronales para tareas de minería de datos,» Santa Fe de la Vera Cruz, 2014.
- [50] F. Rosenblatt, «The perceptron, a perceiving and recognizing automaton Project Para,» *Cornell Aeronautical Laboratory*, 1957.
- [51] J. A.-T. Barrera, «Redes Neuronales,» 2016.
- [52] L. C. Machaca Apaza, «Estimación de la evapotranspiración de referencia utilizando modelos de redes neuronales artificiales en función de elementos climáticos en la cuenca del río Huancané,» 2016.
- [53] L. Zapata, E. , «Entrenamiento de una Red Neuronal Hardware desde MatLab (Hardware in the Loo,» 2017.
- [54] E. Serna, «Desarrollo e innovación en ingeniería,» 2018.
- [55] P. S. Palash, Goyal y K. Jain , «Deep learning for natural language processing,» pp. 138-143, 2018.
- [56] D. R. Lillo y C. Salas, «Efecto de la retroalimentación del error en el aprendizaje y emociones de estudiantes de enseñanza básica.,» *Paideia*, n° 56, pp. 11-42, 2015.
- [57] E. Kuriscak, P. Marsalek y J. Stroffek, «Biological context of Hebb learning in artificial neural networks,» *Neurocomputing*, pp. 27 - 35, 2015.
- [58] D. A. Gavilanes Proaño, «Sistema de monitoreo apícola mediante el uso de redes neuronales artificiales para identificar la variación de población,» 2020.
- [59] A. Padrón y J. Pérez, «Dinámica de una red neuronal artificial con forzamiento periódico.,» 2018.
- [60] E. Muñoz, «Aplicación de redes neuronales para predecir el éxito de la compra de deuda a una entidad financiera.,» *Revista Espacios*, vol. 40, n° 20, 2019.
- [61] M. E. Aruneri Ponce y E. K. Espinal, «Perceptrón memoria de correlación multicapa para predecir la violencia sistemática contra la mujer por su pareja en la Sierra del Perú.,» 2018.
- [62] S. Dutta, «Learning and verification of feedback control systems using feedforward neural networks.,» *IFAC-PapersOnLine*, vol. 51, n° 16, pp. 151-156, 2018.