



POSGRADOS

MAESTRÍA EN

ELECTRÓNICA Y AUTOMATIZACIÓN

RPC-SO-19-No.277-2018

OPCIÓN DE
TITULACIÓN:

PROYECTOS DE DESARROLLO

TEMA:

DESARROLLO DE UN SISTEMA DE PREDICCIÓN DE
FALLA DE COJINETE DE TURBINA HIDRÁULICA BASADO
EN REDES NEURONALES RECURRENTE CON LSTM

AUTORES:

ESTEBAN FRANCISCO REINO CHÉRREZ
JHONNATAN RODRIGO COCHANCELA ARAUJO

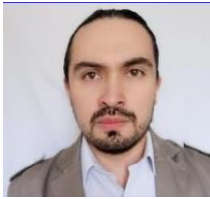
DIRECTOR:

CHRISTIAN RAÚL SALAMEA PALACIOS

CUENCA - ECUADOR

2022

Autores:



Esteban Francisco Reino Chérrez

Ingeniero Electrónico

Candidato a Magíster en Electrónica y Automatización,
Mención en Informática Industrial por la Universidad
Politécnica Salesiana - Sede Cuenca.

ereino@est.ups.edu.ec



Jhonnatan Rodrigo Cochancela Araujo

Ingeniero Eléctrico

Candidato a Magíster en Electrónica y Automatización,
Mención en Informática Industrial por la Universidad
Politécnica Salesiana - Sede Cuenca.

jcochancelaa@est.ups.edu.ec

Dirigido por:



Christian Raúl Salamea Palacios

Ingeniero Electrónico.

Máster en Diseño, Gestión y Dirección de Proyectos.

Doctor en el Programa Oficial de Doctorado en
Ingeniería de Sistemas Electrónicos por la
Universidad Politécnica de Madrid.

csalamea@ups.edu.ec

Todos los derechos reservados.

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra para fines comerciales, sin contar con autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual. Se permite la libre difusión de este texto con fines académicos investigativos por cualquier medio, con la debida notificación a los autores.

DERECHOS RESERVADOS

©2022 Universidad Politécnica Salesiana.

CUENCA – ECUADOR – SUDAMÉRICA

REINO CHÉRREZ ESTEBAN FRANCISCO

COCHANCELA ARAUJO JHONNATAN RODRIGO

***DESARROLLO DE UN SISTEMA DE PREDICCIÓN DE FALLA DE COJINETE
DE TURBINA HIDRÁULICA BASADO EN REDES NEURONALES
RECURRENTE CON LSTM***

Índice general

Índice de Figuras	V
Índice de Tablas	VIII
Resumen	IX
Abstract	IX
1. Introducción	1
1.1. Descripción general del problema.....	2
1.2. Objetivos.....	4
1.2.1. Objetivo general.....	4
1.2.2. Objetivos específicos.....	4
1.3. Contribuciones.....	4
1.4. Organización del manuscrito.....	5
2. Antecedentes y fundamentación teórica	6
2.1. Antecedentes.....	7
2.2. Fundamentación teórica.....	12
2.2.1. Generación de energía hidroeléctrica.....	12
2.2.2. Mantenimiento predictivo o basado en condición (CBM)	16
2.2.3. Modelos matemáticos para la predicción de series temporales.....	20
2.2.4. Regresión lineal simple.....	21
2.2.5. Redes neuronales.....	28
2.2.6. Redes neuronales recurrentes (RNN's).....	28
2.2.7. Long short-term memory (LSTM).....	29

3. Desarrollo del sistema	32
3.1. Metodología del desarrollo.....	33
3.2. Método de aprendizaje.....	34
3.3. Estructura y validación de la base de datos	35
3.3.1. Limpieza de datos.....	36
3.3.2. Selección de datos y parámetros.....	36
3.3.3. Normalización de datos	40
3.4. Diseño del sistema de pronóstico mediante regresión.....	40
3.4.1. Modelo de regresión lineal con datos diarios	41
3.4.2. Modelo de regresión lineal con datos semanales.....	42
3.4.3. Modelo de regresión lineal con datos mensuales.....	44
3.5. Diseño de sistema de pronóstico mediante RNN con LSTM	45
3.5.1. Optimización de la red neuronal.....	47
3.5.2. Balanceado de clases normales y de falla.....	54
3.5.3. Modelado LSTM con envolvente diaria.....	55
3.5.4. Modelado LSTM con envolvente semanal.....	58
3.5.5. Modelado LSTM con envolvente mensual	61
4. Análisis de resultados	63
4.1. Evaluación de modelos de regresión lineal	64
4.2. Evaluación de modelos LSTM	64
4.2.1. Evaluación de red neuronal LSTM con envolvente diaria	66
4.2.2. Evaluación de red neuronal LSTM con envolvente semanal	69
4.2.3. Evaluación de red neuronal LSTM con envolvente mensual	72
5. Conclusiones y trabajo futuro	77
Glosario	87

Índice de Figuras

1.1. Esquema Generador – Turbina: a) Turbina hidráulica, b) Cojinete de turbina hidráulica, c) Eje, d) Generador eléctrico.	2
1.2. Comportamiento de temperatura axial de cojinete de turbina hidráulica del grupo 6 de Saymirín III-IV.....	2
1.3. Cojinete de turbina hidráulica del grupo generador 6.	3
2.1. Visualización de grupo generador 6 de Saymirín III-IV.....	10
2.2. Proceso de conversión de energía [R. Ortiz Flórez, 2011].	13
2.3. Grupo generador unidad 6, Saymirín III-IV, ELECAUSTRO.	14
2.4. Esquema unifilar simplificado de una mini central [J. F. Sanz Osorio, 2016].	16
2.5. a) ausencia de relación (independencia), b) asociación lineal positiva (variación en el mismo sentido), c) asociación lineal negativa (variación en sentido contrario), d) fuerte asociación no lineal.....	22
2.6. Residuos como distancias verticales desde la ecuación estimada de regresión [Weisberg, 2014].....	24
2.7. Ilustración esquemática de la naturaleza iterativa del proceso de regresión [Chatterjee and Hadi, 2012]	27
2.8. Neurona biológica genérica versus representación clásica de neurona artificial [Kattan et al., 2011]	28
2.9. Estructura de una RNN [Ye et al., 2019]	29
2.10. Estructura de LSTM [Mike Bernico, 2018].....	30
3.1. Metodología propuesta mediante regresores a desarrollarse.....	33
3.2. Metodología propuesta mediante RNN con LSTM a desarrollarse	34
3.3. Parte del archivo principal csv.	35
3.4. Histograma de distribución de los datos.	36
3.5. Medición histórica de temperatura axial 2015 a 2021.	37

3.6. Mediciones históricas de temperatura en estado normal (azul) y en falla (rojo) 2015 a 2021.....	38
3.7. Mediciones históricas de temperatura en estado normal 2015 a 2021.....	39
3.8. Mediciones históricas de temperatura en estado de falla 2015 a 2021.....	39
3.9. Modelo de regresión lineal con datos diarios	42
3.10. Modelo de regresión lineal con datos semanales	43
3.11. Modelo de regresión lineal con datos mensuales.....	44
3.12. Modelo inicial de la red neuronal con LSTM	45
3.13. Función de pérdida del modelo inicial.....	46
3.14. Función de pérdida con optimizador Adam	49
3.15. Visualización mediante mapa de calor de los mejores parámetros	51
3.16. Función de pérdida con optimización de capas neuronales y algoritmo Adam.....	52
3.17. Histograma de estados del generador	54
3.18. Función de pérdidas para envolvente diaria con sobremuestreo	56
3.19. Función de pérdidas para envolvente diaria con submuestreo .	57
3.20. Función de pérdidas para envolvente diaria con técnicas combinadas	58
3.21. Función de pérdidas para envolvente semanal con sobremuestreo	59
3.22. Función de pérdidas para envolvente semanal con Cluster Centroids	60
3.23. Función de pérdidas para envolvente semanal con técnicas combinadas	61
3.24. Función de pérdidas para envolvente mensual.....	62
4.1. Envolvente diaria con entrenamiento, validación, prueba y predicción.....	66
4.2. Serie temporal de temperatura de cojinete diaria, color azul corresponde a datos reales, color anaranjado corresponde a la predicción del modelo RNN-LSTM	67
4.3. Aplicación del algoritmo free-running en envolvente diaria	68
4.4. Envolvente semanal con entrenamiento, validación, prueba y predicción.....	70
4.5. Serie temporal de temperatura de cojinete semanal, color azul corresponde a datos reales, color anaranjado corresponde a la predicción del modelo RNN-LSTM	71
4.6. Aplicación del algoritmo free-running en envolvente semanal .	72

4.7. Envolverte mensual con entrenamiento, validación, prueba y predicción.....	73
4.8. Serie temporal de temperatura de cojinete mensual, color azul corresponde a datos reales, color anaranjado corresponde a la predicción del modelo RNN-LSTM.....	74
4.9. Aplicación del algoritmo free-running en envolverte mensual .	75

Índice de Tablas

2.1. Características de las unidades de generación de la central Saymirín III-IV.....	9
2.2. Bases de Datos.....	12
3.1. Resultados batch vs épocas con Adam.....	48
3.2. Resultados batch vs épocas con RMSPROP.....	48
3.3. Resultados batch vs épocas con Adagrad.....	48
3.4. Resultados Neuronas capa Dense vs Neuronas LSTM para RMSE.....	50
3.5. Resultados Neuronas capa Dense vs Neuronas LSTM para coeficiente R^2	50
3.6. Resultados aplicación dropout capa Dense vs capa LSTM para error RMSE.....	53
3.7. Resultados aplicación dropout capa Dense vs capa LSTM para coeficiente R^2	53
3.8. Parámetros de red neuronal inicial y optimizada.....	54
4.1. Resumen de métricas de evaluación para regresión.....	64
4.2. Comparación de métricas de evaluación para envolvente diaria.....	65
4.3. Comparación de métricas de evaluación para envolvente semanal.....	65
4.4. Comparación de métricas de evaluación para envolvente mensual.....	65
4.5. Valores de temperatura de cojinete de turbina, caudal y potencia de la unidad 6 de Saymirin III-IV presentados el 03-04-2021.....	69
4.6. Métricas de evaluación para los modelos de pronóstico desarrollados.....	76

Resumen

Actualmente los procesos industriales hacen uso de los últimos avances tecnológicos y las tendencias de mantenimiento se inclinan por los enfoques de predicción de eventos (PdM). Este trabajo presenta un sistema de predicción de falla de un cojinete de turbina hidráulica mediante la aplicación de dos técnicas de machine learning: modelos de regresión y modelos de redes neuronales recurrentes LSTM (Long short-term memory). Mediante histogramas se visualizó la cantidad elevada de información, se realizó un proceso de validación de la base de datos en el cual se realiza una limpieza, selección y normalización de la información, con la finalidad de facilitar el procesamiento y análisis de los mismos cuando ingresen al sistema de predicción. Se presentan modelos mediante regresión lineal, de estos se establecen las mejoras relativas entre un modelo y otro, determinando el mejor modelo para la predicción mediante esta técnica.

Se realizaron pruebas para definir la red neuronal optimizada: número de neuronas LSTM, neuronas dense, el dropout LSTM/dense, la elección del mejor algoritmo optimizador, el número de épocas y batch, teniendo como métrica de evaluación el RMSE y el R^2 , para posteriormente realizar un balanceo de información, puesto que se disponía de datos mayoritariamente en estado normal y esto podría afectar a la capacidad de generalización de los algoritmos, para este fin se aplicaron técnicas de sobremuestreo (SMOTE), submuestreo (Cluster Centroids) y una combinación de estas, aplicadas a los modelos LSTM de envolvente diaria, semanal y mensual. Se presenta un análisis real de las fechas de pronóstico que entrega el sistema con las dos técnicas, siendo el mejor modelo en base a las métricas de evaluación el logrado con la red neuronal LSTM de envolvente diaria y técnica de sobremuestreo SMOTE.

Abstract

Currently, industrial processes make use of the latest technological advances and maintenance trends lean towards event prediction approaches (PdM). This work presents a failure prediction system for a hydraulic turbine bearing by applying two machine learning techniques: regression models and models with recurrent neural networks LSTM (Long short-term memory). By means of histograms the high amount of information was visualized, database validation process was carried out in which a cleaning, selection and normalization of the information is carried out, in order to facilitate their processing and analysis when they enter the prediction system. The models are presented by linear regression, from these the relative improvements between one model and another are established, determining the best model for prediction using this technique. Tests were performed to define the optimized neural network: number of LSTM neurons, dense neurons, LSTM/dense dropout, election of the best optimizing algorithm, batch and epoch number, having as evaluation metric the RMSE and R^2 , to later carry out a balancing of information, due to the fact that the data was mostly available in a normal state and this could affect the algorithms for generalizing the information, for this purpose, oversampling techniques (SMOTE), subsampling (Cluster Centroids) and a combination of these applied to the daily, weekly and monthly envelope LSTM models were applied. A real analysis of the forecast dates provided by the system with the two techniques is presented, the best model based on the evaluation metrics being the one achieved with the LSTM daily enveloping neural network and the SMOTE oversampling technique.

Capítulo 1

Introducción

En una central hidroeléctrica el mantenimiento preventivo bien organizado puede disminuir considerablemente la probabilidad de ocurrencia de fallas que pueden producir una parada de emergencia de un generador, excepto las fallas catastróficas. Este tipo de mantenimiento brinda grandes beneficios: optimización en el funcionamiento de las maquinarias, ahorro de energía, disminución de costos administrativos, y aumentar la confiabilidad de los dispositivos, al reducirse el porcentaje de averías, tiempo entre fallas, tiempo de paradas y tiempos de reparación. Hoy en día los procesos industriales están siendo fuertemente impulsados por los últimos avances tecnológicos, las tendencias de mantenimiento se inclinan por los enfoques de predicción de eventos (PdM), actualmente no existe una técnica predominante para obtener la predicción más precisa, pero hay muchos enfoques como la regresión lineal, la cual investiga y modela la relación entre variables, o el uso de redes neuronales que consiste en utilizar un conjunto de algoritmos estadísticos y matemáticos para realizar tareas como modelado predictivo, el cual su objetivo final es mejorar el aprendizaje de tal manera que se vuelva automático. El enfoque de este proyecto está orientado al uso de redes neuronales recurrentes (RNN), las cuales pueden aprender características de datos secuenciales. Estas técnicas han demostrado ser efectivas en el área de investigación del diagnóstico de fallas en máquinas [Zhu et al., 2019]. El funcionamiento de las RNN se ve optimizado cuando se les utiliza en conjunto con una arquitectura LSTM, las cuales amplían su memoria para aprender de experiencias importantes que han pasado hace mucho tiempo, además de que pueden capturar y descubrir características significativas para la predicción del estado mecánico (PMS), esto gracias a la capacidad de regulación de datos históricos [Chen and Liu, 2017].

1.1. Descripción general del problema

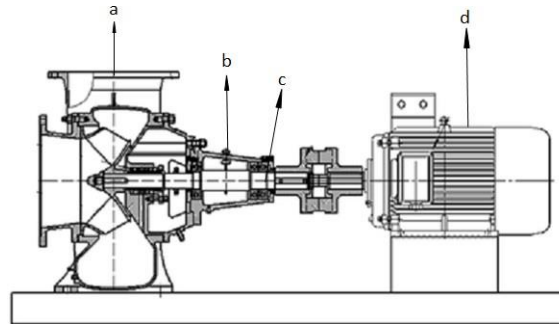


Figura 1.1: Esquema Generador – Turbina: a) Turbina hidráulica, b) Cojinete de turbina hidráulica, c) Eje, d) Generador eléctrico.

Un cojinete de una turbina hidráulica como se aprecia en la Figura 1.1, no hace trabajo útil para la máquina, sino soporta cargas muertas y disipativas [Espejo and Hernández, 2017]. Este elemento tiene una función vital dentro del sistema de un grupo hidroeléctrico, pues es un componente del grupo generador destinado a soportar el giro del eje de la turbina y encargado de suministrar una superficie lubricante para mejorar la fricción del giro.

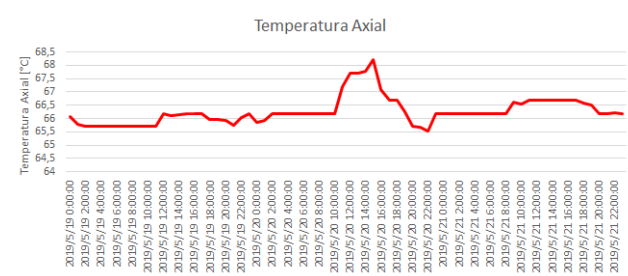


Figura 1.2: Comportamiento de temperatura axial de cojinete de turbina hidráulica del grupo 6 de Saymirín III-IV.

Con el transcurso de los años se ha identificado un problema de sobrecalentamiento en el cojinete de turbina de la unidad de generación 6 de la central Saymirín III-IV, como se puede apreciar en la Figura 1.2, lo que ha obligado que se deba realizar paradas de emergencia debido a fallas de este elemento, provocando pérdidas de producción, aumento de inactividad del

grupo generador, atrasos en planes de mantenimiento preventivos; siendo la alta confiabilidad y disponibilidad con bajos costos de ciclo de vida las metas generales de todos los programas de mantenimiento [Arno et al., 2015].



Figura 1.3: Cojinete de turbina hidráulica del grupo generador 6.

Por esta razón es necesario tener un sistema que tenga la capacidad de predecir averías, pudiendo mejorar con ello los planes de mantenimiento preventivo. Dadas estas condiciones, será importante que el sistema posea características de memoria y de aprendizaje de máquina a partir de los datos, y por ello se han considerado a los sistemas inteligentes automáticos como los idóneos para este objetivo. Actualmente las redes neuronales artificiales se utilizan en aplicaciones de ingeniería de procesos para la predicción y pronóstico de fallas, donde los métodos matemáticos tradicionales no muestran completamente el comportamiento representado por observaciones y resultados experimentales o datos de operación de plantas. Estas aplicaciones son importantes en el área de producción energética y control automático de procesos. Así, las redes neuronales han sido extensamente aplicadas a sistemas de pronóstico y clasificación lineales y no lineales [Calderón et al., 2008]. Entre algunos ejemplos de la aplicación de las redes neuronales en la ingeniería de predicción de fallas podemos mencionar a [Yuhai et al., 2018], donde se presenta un método de predicción y diagnóstico de fallas basado en las señales de vibraciones con la combinación de técnicas de DBN (red de creencia profunda) y LSTM, con resultados que demuestran que este método reduce la dificultad del preprocesamiento de datos y aumenta la precisión de la predicción de fallas. Otras técnicas aplicadas a la creación de

modelos de predicción de falla es la minería de datos; en [Ravi et al., 2019], se identifican las técnicas y precisión de algoritmos comunes de minería de datos que se implementan en estudios relacionados con varios tipos de fallas de transformadores, además de comparar estas técnicas con el algoritmo de una red neuronal destacándose esta al ser la adoptada por la mayoría de investigadores debido a su precisión comparada con otros algoritmos de predicción como árbol de decisión (DT), máquina de soporte vectorial (SVM) y naive bayes (NB).

1.2. Objetivos

1.2.1. Objetivo general

Desarrollar un sistema de predicción de falla de cojinete de turbina hidráulica basado en redes neuronales recurrentes con LSTM

1.2.2. Objetivos específicos

- Procesar la información de la base de datos, estructurada con las variables implicadas en la medición de la temperatura del cojinete, mediante la aplicación de técnicas de análisis de datos, minimizando la redundancia de datos y eliminando los valores que se encuentran fuera del rango de este estudio.
- Implementar el sistema de predicción de falla del cojinete utilizando la técnica de regresión lineal obteniendo así la métrica de base del sistema inteligente.
- Implementar el sistema de predicción de falla del cojinete utilizando redes neuronales recurrentes con LSTM, mediante el uso de herramientas informáticas especializadas verificando la utilidad de la información basada en la definición de la probabilidad de hechos futuros.
- Comparar la eficiencia de entre las dos técnicas por medio de sus curvas de predicción determinando la más adecuada, considerando las variables de entrada.

1.3. Contribuciones

Las contribuciones de este trabajo son las siguientes:

1. El desarrollo de un sistema de pronóstico de fallas en el cojinete de turbina hidráulica de un grupo generador para la mejora de los planes de mantenimiento preventivo, contribuyendo a una reducción de fallos y averías, reducción del número de paradas no programadas, y al aumento de la disponibilidad del grupo generador.
2. La mejora de los planes de mantenimiento preventivo para la prolongación de la vida útil de los elementos mecánicos, así como la reducción de inversión en repuestos al extender el ciclo de vida del activo.
3. Comparativa del grado de eficiencia de pronóstico entre las técnicas de regresión lineal y la red neuronal recurrente con LSTM determinando la más adecuada para este trabajo.
4. Este proyecto contribuye a un futuro desarrollo de más sistemas de pronóstico de fallas con aplicación a otros sistemas de el resto de centrales hidroeléctricas de la empresa ELECAUSTRO.

1.4. Organización del manuscrito

El presente documento se encuentra integrado de la siguiente manera: El capítulo 2 se encuentra constituido por los antecedentes de predicción y pronóstico de fallas dentro de la industria, para cojinetes y elementos rotativos; características de la central Saymirín III-IV e información de las BD del Sistema SCADA. Además contiene la base teórica para este trabajo, como son los temas de generación de energía hidroeléctrica, el mantenimiento predictivo y los modelos matemáticos utilizados para el pronóstico de series temporales.

El capítulo 3 se describen los procedimientos seguidos para el desarrollo del sistema de pronóstico de fallos, tanto con los modelos de regresión, estableciendo un baseline para el sistema y el procedimiento de diseño de la red LSTM: validación de BD, optimización y balanceo de datos.

El capítulo 4 contiene el análisis de resultados de las dos técnicas de pronóstico aplicadas al proyecto.

El capítulo 5 contiene las conclusiones del proyecto y trabajos a futuro que se podrán desarrollar para el pronóstico de falla de diferentes sistemas de las centrales de ELECAUSTRO.

Capítulo 2

Antecedentes y fundamentación teórica

El capítulo se encuentra constituido por los antecedentes de pronóstico de fallas dentro de la industria, para cojinetes y elementos rotativos; características de la central Saymirín III-IV e información de las BD del Sistema SCADA. Además contiene la base teórica para este trabajo, como son los temas de generación de energía hidroeléctrica, el mantenimiento predictivo y los modelos matemáticos utilizados para el pronóstico de series temporales.

2.1. Antecedentes

Los sistemas de predicción de fallas dentro de la industria de generación de energía se han convertido en un punto clave al momento de asegurar un óptimo funcionamiento de los activos y la confiabilidad de los sistemas. La predicción de las tendencias de degradación de los rodamientos es importante para la estimación de la vida útil restante (RUL) [Baumgart et al., 2015]. Parte importante dentro de este grupo industrial son los elementos rotativos es por esto que se han desarrollado varios estudios y metodologías para llegar a anticiparse a una posible falla. Existen muchos trabajos que abordan los pronósticos de fallas y gestión de salud (PHM) de cojinetes y elementos rotativos, en los que se aplican varias técnicas como en [Deutsch and He, 2017], que presenta un enfoque basado en deep learning para la predicción del RUL con el uso de técnicas de big data y se comparan con métodos de PHM existentes, logrando resultados favorables. En cuanto a realizar una predicción con modelos estadísticos se puede revisar [Wang and Tsui, 2017], donde propone un modelo de señales de degradación de rodamientos y detección temprana en el elemento, y luego el modelo para la predicción del RUL logrando una superioridad al modelo estadístico propuesto previamente en [Taylor et al., 2007]. Unas de las variables que más información pueden entregar para realizar una predicción de fallo de un cojinete es el análisis de vibraciones, como la presentada en [Shuqing, 2006], en la que se aplican técnicas de transformada rápida de fourier (FFT), transformada de fourier de corta duración (STFT), redes neuronales y análisis de ondas aplicadas en una turbina hidroeléctrica, de igual manera [Chang, 2019, Siegel et al., 2011], tratan estudios basados en las vibraciones en las que se extraen características de tiempo y frecuencia para construir modelos de regresión para la predicción de fallas en rodamientos de un motor de inducción y cojinetes de helicópteros [Lt et al., 2016] expone un método integrado para detectar los defectos incipientes de los rodamientos mediante análisis de vibraciones y extracción de características, utilizando descomposición de wavelets package de 3 capas (WPD) y se presentan los resultados en un análisis comparativo con las características estadísticas clásicas del dominio del tiempo.

Otra técnica empleada para pronósticos de degradación de rodamientos es utilizar un algoritmo de vectores de regresión (SVR) como se plantea en [Loutas et al., 2013], el cual se basa en buscar un hiperplano que modele la tendencia de los datos de entrenamiento y con esta información predice datos futuros del RUL, utilizando registros de formas de onda de aceleración. Otros métodos como [Marble and Morton, 2006] presenta

el estudio de pronóstico de degradación basado en un modelado de los desprendimientos del material metálico de cojinetes, calculando la trayectoria de crecimiento de alimañas y el tiempo hasta la falla en función de las condiciones de funcionamiento, con retroalimentación para autoajustarse y lograr una reducción de la incertidumbre de la predicción. Por otra parte [Boukra and Lebaroud, 2014] presenta un interesante estudio donde presenta nuevas características de base para el estudio para predicción, como es el algoritmo de optimización por enjambre de partículas (PSO) para identificar características de pronóstico avanzado y el Modelo Oculto de Semi Markov (HSMM) para predicción del RUL. Otros métodos como el presentado en [Wang et al., 2020] se han enfocado en utilizar sólo datos de estado normal de la máquina y sin un conocimiento de la estructura del equipo, presentando un modelo denominado WPD-DKPCA basado en machine learning, logrando resultados de predicción confiables y precisos. En [Caesarendra et al., 2011] se propone una aplicación de modelos de máquina de vectores de relevancia (RVM), regresión logística (LR) y modelos de media móvil autorregresiva y de heterocedasticidad condicional autorregresiva generalizada (ARMA / GARCH) para evaluar la degradación de fallas en base a datos de simulación de rodamientos de ejecución a falla, luego se compara el resultado con el modelo de regresión de dempster-shafer (DSR), obteniendo un modelo válido de pronóstico de degradación de rodamientos.

La empresa Electro Generadora del Austro, ELECAUSTRO S.A. cuenta con varias centrales de generación, entre las cuales está la central hidroeléctrica Saymirín III-IV, construida en 1995 con dos grupos generadores de 4 MW cada uno y turbinas tipo Francis.

En la Tabla 2.1, se presenta las características de las unidades de generación de la central Saymirín III-IV.

Tabla 2.1: Características de las unidades de generación de la central Saymirín III-IV.

Componente	Característica	Unidades 5 y 6
Generador	Marca	BBC
	Tipo	HC17/23
	Enfriamiento	Aire Forzado
	N de Polos	12
	Factor de potencia	0,8
	Potencia [KVA]	5700
	Tensión [KV]	2,4
	Frecuencia [Hz]	60
	Conexión	Y con transferencia
	Aislamiento	F
Excitación	Tipo	Diodos Rotativos
	Corriente	395 - 460
	Tensión	130 DC
Turbina	Marca	BELL
	Tipo	Francis
	Potencia [KW]	4000
	Velocidad [RPM]	900
	Embalamiento [RPM]	1360
	Caudal [M3/Seg.]	2,143
	Caida neta (placa) [M]	212
Transformador	Tipo	En aceite
	Enfriamiento	ONAF
	Capacidad [MVA]	10
	Voltaje [KV] 1rio:	69
	Voltaje [KV] 2rio:	2,4
	Conexión 1rio:	Y a tierra
	Conexión 2rio:	D
	Fabricante	BBC

Desde sus inicios ELECAUSTRO S.A. se ha caracterizado por ser una empresa en constante crecimiento y de proyección siempre hacia el futuro, generando nuevos proyectos y procesos de modernización. La central cuenta con el sistema SCADA (supervisión, control y adquisición de datos) OASyS de Telvent, en el cual la gestión de bases de datos es un componente central y está organizado en dos partes fundamentales, una optimizada para el manejo de datos en tiempo real y la otra para la administración de datos históricos. Dentro de los beneficios que brinda el sistema SCADA se encuentra el contar con una amplia base de datos de señales eléctricas y mecánicas que podrían ser de utilidad en la implementación de sistemas de aprendizaje automático.

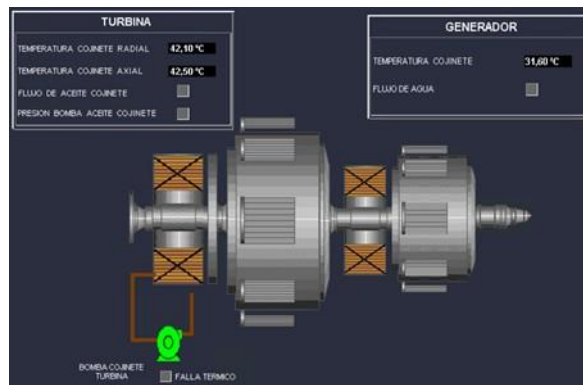


Figura 2.1: Visualización de grupo generador 6 de Saymirín III-IV

El sistema de gestión de bases de datos ofrece todas las interfaces necesarias para mover información entre los componentes. Así, permite estructurar la información, almacenar los datos adquiridos y proporcionar un entorno para el procesamiento de datos, control de dispositivos y supervisión interna de los procesos del sistema.

La BD de tiempo real está diseñada especialmente para el procesamiento de datos instantáneos. Los datos de configuración y los datos de tiempo real acerca de dispositivos están contenidos en la tabla. El número de registros relativos a información del sistema productivo permanece constante, por lo que el espacio ocupado siempre es el mismo, siendo esta circunstancia vital para evitar el desbordamiento de la memoria.

El sistema hace peticiones constantemente a las fuentes de datos, sobrescribiendo los datos con los nuevos valores obtenidos, por otra parte, la base de datos de históricos del sistema OASyS DNA, está basada en Microsoft SQL server standard edition. La capacidad de almacenamiento de dicha base de datos sólo está limitada por el hardware y el sistema operativo. El administrador del sistema tiene la capacidad de definir los datos analógicos que desea almacenar mediante definiciones que pueden ser editadas y visualizadas a través de la herramienta ADE (Advance Database Editor), en base a esta definición se recopila los datos históricos a intervalos periódicos.

Adicionalmente, y según intervalos de una hora, se ejecuta una tarea que calcula para cada uno de los períodos:

- Valor máximo en intervalo de tiempo.
- Instante en que se produjo el máximo.
- Valor mínimo en intervalo de tiempo.
- Instante en que se produjo el mínimo.
- Valor medio con la calidad del dato.
- Valor instantáneo con la calidad del dato.

La BD de históricos está compuesta de una serie de bases de datos, cada una con sus tablas y vistas. Las vistas son tablas de sólo lectura que muestran los campos más utilizados. Los datos son almacenados en las siguientes bases de datos:

Tabla 2.2: Bases de Datos

Base de Datos	Tablas	Vistas	Descripción
Accum	Hour Day Month Year	HourView DayView MonthView YearView	Datos acumulados de contadores
Archive	Catalog Device DumpSchedule Rearchive Schedule ValidDeviceTypes	CatalogView DeviceView RearchiveView ScheduleView	Datos de configuración y archivado
CommStats	RemPeriodStats ConnPeriodStats	RemPeriodStatsView EditRemPeriodView ConnStatsView EditConnPeriodView	Estadísticas de comunicaciones
Event	Summary	EventView SummaryView	Mensajes: eventos y alarmas
TimeSeries	Hour Day Month Year Collect Tag Disturbevent Disturbanalog	HourView DayView MonthView YearView	Datos numéricos de Realtime
Xosapp	Notepad	NotepadView	Otras aplicaciones XOS

La variable bajo estudio “temperatura de cojinete de turbina” se encuentra en la base de datos TimeSeries y consta con registros desde el 01 de enero del 2015.

2.2. Fundamentación teórica

2.2.1. Generación de energía hidroeléctrica

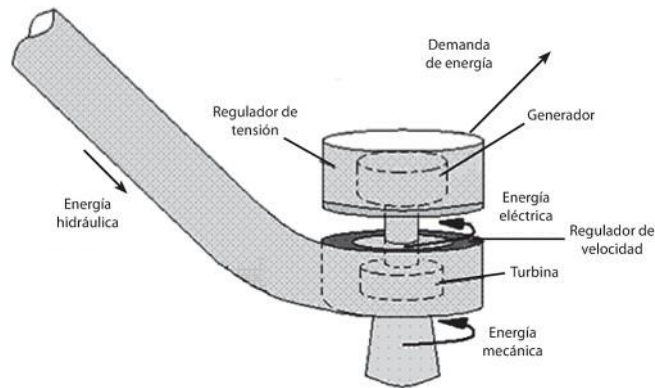


Figura 2.2: Proceso de conversión de energía [R. Ortiz Flórez, 2011].

Una central hidroeléctrica es una instalación cuya misión es convertir la energía potencial y cinética del agua en energía eléctrica disponible. Se encuentra formada por varios elementos que intervienen en esta transformación como:

- Elementos de retención y almacenaje de agua.
- Elementos de conducción de agua tales como canales, tuberías.
- Elementos de apertura y cierre del paso de agua: compuertas, válvulas.
- Equipamiento hidráulico: turbina, multiplicador, rejas
- Equipamiento eléctrico: generador, transformador, línea eléctrica.
- Equipamiento de control y protecciones: interruptores, seccionadores, etc.
- Equipamiento Auxiliar: baterías de corriente continua, tomas de fuerza, etc.

Un grupo generador hidroeléctrico consta de los elementos eléctricos y mecánicos que permiten que se produzca esta transformación de energía cinética del agua hacia energía mecánica de rotación, permitiendo que las turbinas giren a alta velocidad, produciendo el cambio de energía mecánica a energía eléctrica. Los elementos de un grupo generador pueden apreciarse en la Figura 2.3.



Figura 2.3: Grupo generador unidad 6, Saymirín III-IV, ELECAUSTRO.

a) Turbina

La turbina es el equipo en el cual se transforma la energía hidráulica en mecánica, está acoplada directamente al generador y en conjunto atienden la demanda de energía eléctrica. Estas se dividen en dos grupos como son las turbinas de acción y reacción. En las de acción el caudal cambia solamente de dirección más no de aceleración, mientras que las de reacción el caudal entra a presión y en los conductos móviles del rodete cambia de dirección y aceleración [R. Ortiz Flórez, 2011].

b) Cojinete de turbina y de generador

Su función es la soportar los ejes de transmisión de potencia, además de la reducción del rozamiento que se produce entre los ejes, por lo que son encargados de suministrar una superficie de lubricante para evitar el contacto metálico, así logrando reducir la fricción de deslizamiento de los rodamientos.

c) Generador

El generador tiene como misión transformar en energía eléctrica la energía mecánica suministrada por la turbina. Está formado por una parte estática, denominada estator, y una parte móvil, denominada rotor.

d) Grupo oleohidráulico

Su función es la de por medio de sus bombas permitir la circulación desde el tanque de reserva el aceite de lubricación hacia el cojinete de turbina.

e) Excitatriz

Tiene la función de ser una fuente de corriente continua para la alimentación de las bobinas de excitación de los polos inductores, situados en el rotor. De esta forma se consigue generar un campo magnético giratorio y senoidal en el entrehierro [J. F. Sanz Osorio, 2016].

f) Regulador de velocidad

El regulador es un sistema de control que a partir de las variaciones de velocidad envía señales que controlan la apertura o cierre de válvulas provocando variaciones de potencia mecánica de la turbina que compensen las variaciones de potencia eléctrica [Delgado, 1997].

En cuanto a el equipo eléctrico de potencia, este está formado por los elementos necesarios para, una vez obtenida la energía eléctrica en bornes del generador, poder entregarla en el lugar exigido con las características adecuadas [J. F. Sanz Osorio, 2016]. En la Figura 2.4, se observa un esquema unifilar simplificado de una mini central hidráulica con generador síncrono. En ella se pueden observar los principales elementos que la forman:

- Generador.
- Transformador de potencia.
- Interruptores automáticos (identificados por el número 52).
- Seccionadores.
- Elementos de mando y protección.
- Servicios auxiliares.
- Línea eléctrica.

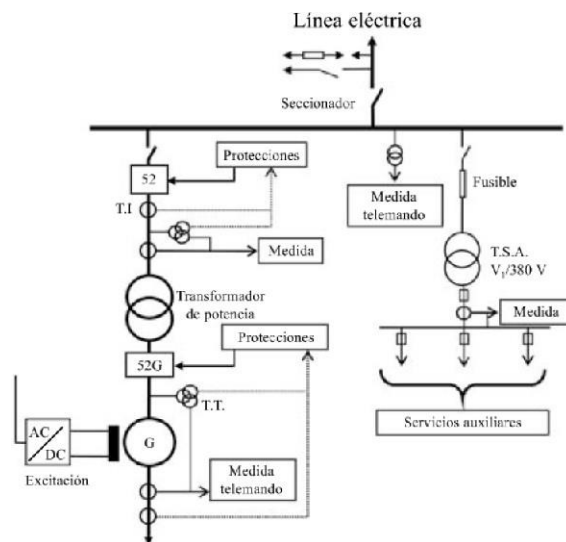


Figura 2.4: Esquema unifilar simplificado de una mini central [J. F. Sanz Osorio, 2016].

2.2.2. Mantenimiento predictivo o basado en condición (CBM)

Dentro de las ramas del mantenimiento existen dos divisiones principales: **Mantenimiento correctivo:** se realiza después del reconocimiento de una avería y que está destinado a poner a un elemento en un estado en que pueda realizar una función requerida.

Mantenimiento preventivo: se realiza a intervalos predeterminados o de acuerdo a criterios establecidos, y que está destinado a reducir la probabilidad de fallo o la degradación del funcionamiento de un elemento [AENOR, 2011].

El mantenimiento predictivo es un conjunto de técnicas instrumentadas de medidas y análisis de variables, implementadas para identificar, en términos de fallas potenciales, la condición operativa de los equipos involucrados en el proceso de la planta, el objetivo principal es optimizar la fiabilidad y disponibilidad de los equipos sin elevar costos de mantenimiento, para ello se aplican técnicas especializadas que ayudan a detectar con anticipación un desperfecto en el equipo, el mal funcionamiento o el cambio de estado de un equipo o máquina durante su operación [Ajuech, 2017].

Un buen programa de mantenimiento predictivo puede generar un ahorro anual del 8 % al 12 %; en el mejor de los casos, dependiendo de las condiciones de equipo y material, es posible llegar a un ahorro del 30 % a 40 % [Poor et al., 2019].

En smart industries, el mantenimiento predictivo es una de las técnicas más utilizadas para abordar este tipo de problema, ya que permite evaluar las condiciones de un equipo específico, con el objetivo de predecir el mantenimiento antes de la falla [Bianchi et al., 2014], estos indicadores pueden ser las vibraciones anómalas, temperaturas elevadas, potencia absorbida, análisis de lubricantes; estos parámetros brindan información sobre un comportamiento fuera de los rangos de operación normales, con los datos que nos brindan se confeccionan los planes de mantenimiento y las intervenciones programadas [Gallará, I. y Pontelli, 2020], con los indicadores mencionados, dentro del mantenimiento basado en condiciones, se pueden utilizar para detectar, identificar y predecir el inicio y la evolución de posibles fallas. Luego, la información se utiliza para programar de manera óptima las actividades de mantenimiento y la logística. Con el desarrollo del aprendizaje automático y el big data, los algoritmos de aprendizaje profundo se convierten en herramientas importantes en la predicción de condiciones debido a sus excelentes capacidades en el procesamiento de datos, extracción de características y modelado [Niu et al., 2018].

Dentro de la literatura que abarcan las LSTM en el mantenimiento predictivo se encuentra [Vildan et al., 2020], que modela la degradación de un motor turbofan para predecir el RUL del conjunto de datos del motor, para pronosticar las necesidades de mantenimiento del motor antes de fallar. Para ello, entrenan dos estructuras de redes neuronales artificiales profundas (LSTM y WaveNet) y se realizan pruebas utilizando los modelos obtenidos como resultado del entrenamiento, con el resultado que la estimación del modelo y los datos reales son muy cercanos entre sí, especialmente en los periodos cercanos al tiempo de mantenimiento, demostrando que el modelo creado puede estimar con precisión los requisitos de mantenimiento y funcionar de acuerdo con su propósito.

Muchos trabajos se han realizado en el campo del pronóstico de fallas de maquinarias con la aplicación de RNN como en [Zhang et al., 2021], que presenta un nuevo método de diagnóstico de fallas para maquinaria rotativa basada en GRU (gated recurrent unit) donde se aprovecha en gran medida de la información de los datos de series temporales y se aprende las características representativas de las señales de vibración, según las características aprendidas, el método propuesto FDGRU (gated recurrent unit recurrent neural network for fault diagnosis) logra el mejor rendimiento

en los conjuntos de datos CWRU (Case Western reserve University) y SPCP (Self - priming centrifugal pump) en comparación con los métodos de deep learning existentes o tradicionales. Además, FDGRU exhibe un sólido desempeño contra el ruido, lo cual es esencial en las condiciones de trabajo extremas, por lo tanto, este trabajo se puede considerar como una guía para reducir la influencia del ruido en el diagnóstico de fallas.

Otro método [Arunthavanathan et al., 2021], se basa en el uso de un modelo de deep learning para el pronóstico de fallas de procesos, el mismo que puede pronosticar características de datos de procesos complejos multivariantes dependientes del tiempo y detectar la condición de anomalía de cada ventana de muestreo. Este sistema es propuesto mediante la interconexión entre CNN (convolutional neural network), LSTM y OC (one class)-SVM. Se utiliza un enfoque de aprendizaje semisupervisado en el modelo CNN LSTM para pronosticar la próxima ventana de muestreo aprendiendo el patrón anterior. Los datos históricos se utilizan para entrenar el modelo CNN LSTM. OC-SVM utiliza un enfoque de aprendizaje no supervisado para definir una condición anormal de los datos sintéticos pronosticados. Los resultados de este trabajo confirman que con un conocimiento limitado de las fallas a partir de los datos sintéticos pronosticados, se puede pronosticar el estado de la falla, además de que el modelo se puede aplicar para predecir el RUL.

Existen interesantes trabajos para la detección de fallas en rodamientos mediante el uso de las RNN como [Xu et al., 2020], que presenta un sistema de diagnóstico de fallas de rodamientos utilizando una red neuronal convolucional multiescala mejorada con un mecanismo de atención de características al que se lo denomina IMS-FACNN, con el que se aborda los bajos rendimientos de los modelos tradicionales basados en CNN en entornos de trabajo inestables y complejos. El mecanismo de atención de características desarrollado mejora la capacidad de generalización del modelo, puede distinguir el tipo de falla, el grado de daño y la ubicación de la falla, tiene un buen desempeño en pruebas de escenarios múltiples, presenta un mejor rendimiento de extrapolación en comparación con WT-CNN (Wavelet transform convolutional neural network), TICNN (Convolution neural networks with training interference), CNN, MS-CNN (Multi-scale convolutional neural network), MC-CNN (multi-scale cascade convolutional neural network) y SVM bajo pruebas de múltiples escenarios. En comparación con los métodos de múltiples escalas existentes, como MS-CNN y MC-CNN, la precisión del IMS-FACNN es superior al 6 % en pruebas de escenarios múltiples. Este modelo IMS-FACNN logra una alta

precisión de diagnóstico de falla para las turbinas eólicas de 750 kW y 3 MW que operan en el mundo real.

[Wang et al., 2019] en la búsqueda de un método efectivo para la identificación de fallas de rodamientos propone y valida un método de búsqueda de arquitectura de redes neuronales basado en el aprendizaje por refuerzo. La metodología propuesta contiene dos componentes: un modelo de controlador y un modelo secundario, el controlador es una red neuronal recurrente (RNN) que genera una serie de acciones, cada acción especifica una opción de diseño para construir los modelos secundarios para el diagnóstico de fallas. Luego, los parámetros del controlador se actualizan utilizando el método de gradiente de políticas de aprendizaje por refuerzo al maximizar la precisión de los modelos secundarios. La arquitectura de red obtenida por el método propuesto puede obtener una alta precisión de clasificación en ambos conjuntos de datos, lo que verifica la universalidad y efectividad de este método en la aplicación práctica, y de gran importancia para la construcción automática de modelos de aprendizaje profundo.

En cuanto a métodos que trabajen con el RUL y las RNN se encuentra [Ding et al., 2021], que propone un método para predecir la RUL de rodamientos basado en la red neuronal convolucional (CNN), se aplica el criterio 3 sigma para eliminar el ruido de los datos originales y eliminar los errores graves. Posteriormente, las características de frecuencia se obtienen a partir de los datos originales mediante la transformada rápida de Fourier (FFT), y la raíz cuadrática media se emplea como métrica de seguimiento. Luego, se aplica muestreo estratificado, que difiere del método tradicional de partición de datos de series de tiempo, se aplica a la partición de datos para conocer completamente las características de los datos. Por lo tanto, la precisión de predicción de DCNN con este método es mejor que la partición tradicional de series de tiempo. Para evitar la pérdida de características, se construye el modelo DCNN sin la capa de agrupación, que consta de tres capas convolucionales y dos capas completamente conectadas. Para la verificación del modelo se muestra que el método de eliminación de ruido basado en el criterio 3 sigma y el muestreo estratificado combinado con DCNN tiene la mayor precisión de predicción y la mayor capacidad de generalización. En los métodos de pronóstico basados en datos, la precisión de la predicción del RUL del rodamiento depende principalmente del rendimiento de los indicadores de salud del rodamiento, que generalmente se fusionan a partir de algunas características estadísticas extraídas de las señales de vibración. Sin embargo, muchos indicadores de salud de rodamientos existentes tienen las siguientes dos deficiencias: (1) muchas características estadísticas no tienen la misma

contribución a la construcción de indicadores de salud ya que los rangos de estas características estadísticas son diferentes; (2) es difícil determinar un umbral de falla ya que los indicadores de salud de diferentes máquinas generalmente son diferentes en un momento de falla, para superar estos inconvenientes.

En [Guo et al., 2017], se propone un indicador de salud basado en redes neuronales recurrentes (RNN-HI) para la predicción RUL de rodamientos. Primeramente, se combina seis características de similitud relacionada con ocho características clásicas de tiempo-frecuencia para formar un conjunto de características original. Luego, con métricas de monotonidad y correlación, las características más sensibles se seleccionan del conjunto de características original. Finalmente, estas características seleccionadas se alimentan a una red neuronal recurrente para construir el RNN-HI. El rendimiento del RNN-HI se verifica mediante dos conjuntos de datos de rodamientos recopilados de experimentos y un campo industrial. Los resultados muestran que el RNN-HI obtiene valores de correlación y monotonidad bastante altos y es beneficioso para soportar la predicción de RUL. Además, se demuestra experimentalmente que el RNN-HI propuesto es capaz de lograr un mejor rendimiento que un método basado en mapas de auto organización.

Hoy en día el mantenimiento industrial es principalmente correctivo (las acciones se toman tras una falla, para corregirla y dejar el equipo productivo) y preventivo, siendo la estrategia predictiva sólo aplicada a situaciones críticas, estas estrategias de mantenimiento no están tomando en cuenta la enorme cantidad de datos que se generan en las industrias y la información emergente disponible con las TIC's, IoT, big data, análisis de datos avanzados, computación en la nube y realidad aumentada [Cachada et al., 2018], para contribuir a mejorar planes de mantenimiento y contribuir a la creación de valor añadido acorde con las necesidades de la empresa [Poor et al., 2019].

2.2.3. Modelos matemáticos para la predicción de series temporales

Un modelo matemático puede ayudar a predecir el comportamiento de una variable en función de un conjunto de otras variables, mientras más relacionadas estén las variables predictoras con la variable a predecir, se obtendrá un mejor resultado de predicción. Entre algunos de los modelos matemáticos para la predicción de series temporales tenemos a el modelo de regresión Lineal, los modelos no lineales GARCH, el cual encuentra la volatilidad promedio a medio plazo mediante una autoregresión que depende

de la suma de perturbaciones rezagadas y de la suma de varianzas del mismo tipo [Francq et al., 2021]. Otro modelo también son los polinomios ortogonales, que a diferencia del polinomio típico basado en potencias crecientes de la variable x , esta clase de polinomios se basan en funciones seno, coseno, o alternativamente, en transformaciones exponenciales de la variable x [Soria Lorente et al., 2014], otro de los modelos de predicción muy utilizado son las redes neuronales. En este trabajo se aplican técnicas de regresión lineal, y una red neuronal LSTM como se verá a continuación.

2.2.4. Regresión lineal simple

El análisis de regresión es un método conceptualmente simple para investigar relaciones funcionales entre variables [Chatterjee and Hadi, 2012], es el estudio de la dependencia, se utiliza para responder preguntas interesantes sobre cómo uno o más predictores influyen en una respuesta. La relación se expresa en forma de una ecuación que conecta la variable respuesta o dependiente a una variable independiente. [Uriel, 1998] menciona que en principio no sabemos si las variables en cuestión están relacionadas o no, o si en caso de haber dependencia es significativa o no. De haber entre ellas una dependencia lineal significativa, podríamos expresar la comprensión (Y) a partir de la presión (X) mediante una recta, y a partir de ella predecir la comprensión que se daría para un determinado nivel de presión. Una forma de determinar si puede existir o no dependencia entre variables, y en caso de haberla, deducir de qué tipo puede ser, es gráficamente representando los pares de valores observados. A dicho gráfico se le llama nube de puntos o diagrama de dispersión, Figura 2.5.

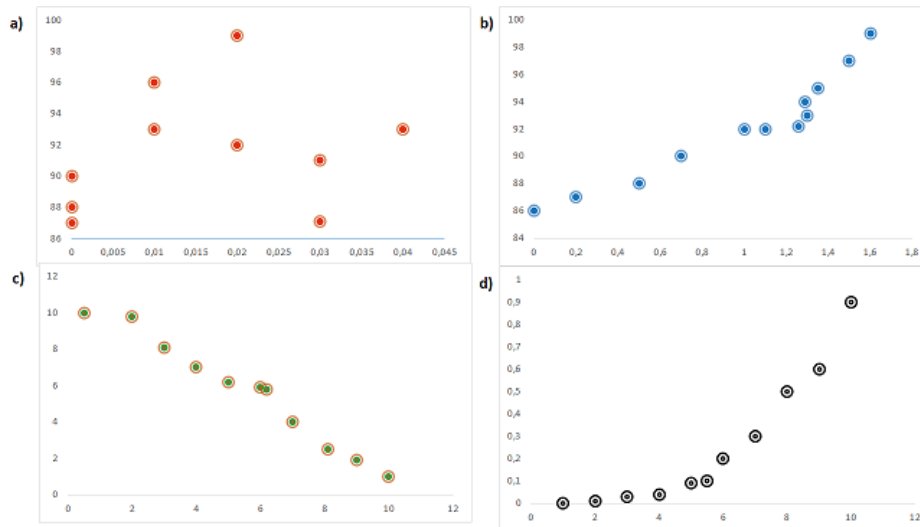


Figura 2.5: a) ausencia de relación (independencia), b) asociación lineal positiva (variación en el mismo sentido), c) asociación lineal negativa (variación en sentido contrario), d) fuerte asociación no lineal.

El método más utilizado en regresión es la regresión lineal, prácticamente todos los demás métodos de regresión se basan en una comprensión de cómo funciona la regresión lineal. Como ocurre con la mayoría de los análisis estadísticos, el objetivo de la regresión es resumir los conjuntos de datos de la manera más simple, útil y elegante posible [Weisberg, 2014].

El modelo de regresión lineal simple está definido por:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (2.1)$$

Donde:

y_i : valor de la variable respuesta en el caso i -ésimo

β_0 : constante o término independiente (intercept)

β_1 : pendiente (slope)

x_i : valor de la variable predictiva en el caso i -ésimo

ε : oscilación aleatoria del caso i -ésimo

Se aprecia en 2.1, que los factores que influyen en la variable respuesta Y pueden dividirse en dos grupos: el que lleva la variable explicativa X y

el segundo que contiene un conjunto amplio de factores no controlados a los que denominaremos como error aleatorio, ε ;, que provoca que la dependencia entre variables dependiente e independiente no sea perfecta, sino esté sujeta a la incertidumbre [Cortés et al., 2015].

Lo deseable en un modelo de regresión es que estos errores aleatorios sean media cero para cualquier valor x de X , es decir:

$$E[\varepsilon/X = x] = E[\varepsilon] = 0 \quad (2.2)$$

Y por lo tanto:

$$E[Y/X = x] = \beta_0 + \beta_1 x + E[\varepsilon/X = x] = \beta_0 + \beta_1 x \quad (2.3)$$

En 2.3 se observa que:

La media de Y , para un valor fijo x , varía linealmente con x . Para un valor x se predice un valor en Y dado por:

$$y = E[Y/X = x] = \beta_0 + \beta_1 x \quad (2.4)$$

Por lo que el modelo de predicción puede expresarse también como:

$$y = \beta_0 + \beta_1 x \quad (2.5)$$

Además de la hipótesis de los errores de que en media han de ser cero, se establecen las siguientes hipótesis:

- La varianza de ε es constante para cualquier valor de x , es decir:

$$V ar[\varepsilon/X = x] = \sigma^2 \quad (2.6)$$

- La distribución de ε es normal, de media 0 y desviación σ .
- Los errores asociados a los valores de Y son independientes unos de otros.

En consecuencia, la distribución de Y para x fijo es normal, con varianza constante σ^2 , y media que varía linealmente con x , dada por $\beta_0 + \beta_1 x$.

Para realizar los pronósticos se deben estimar los parámetros desconocidos β_0 y β_1 , donde β_0 es la ordenada al origen del modelo y β_1 la pendiente, es decir se buscan ambos valores que hagan los mínimos errores de estimación para encontrar la mejor recta que se ajuste a nuestros datos, entonces

aquellos valores de β_0 y β_1 que hagan mínima la suma de los cuadrados de los errores sería:

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2 \quad (2.7)$$

La función de criterio para la obtención de estimadores se basa en residuos, que son las distancias verticales entre la línea ajustada y los valores de y reales, como los ilustrados en la Figura 2.6. Los residuos reflejan la simetría inherente en el rol de la respuesta y el predictor en problemas de regresión [Weisberg, 2014].

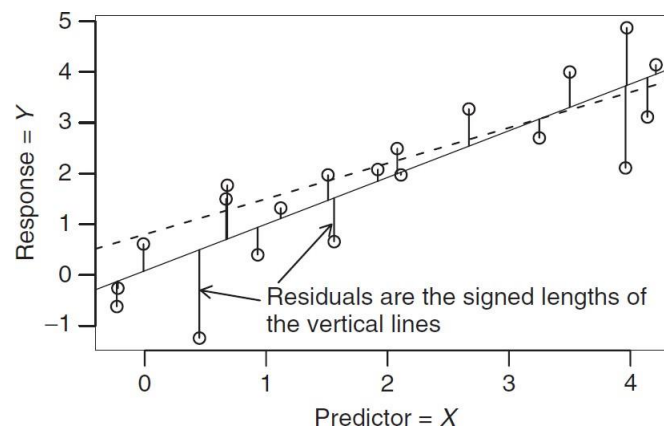


Figura 2.6: Residuos como distancias verticales desde la ecuación estimada de regresión [Weisberg, 2014]

Dado que un residuo representa la cantidad en la que un valor estimado falla para predecir la media de la correspondiente observación aleatoria, entre más grandes son las magnitudes de los residuos, mayor tenderá a ser el efecto de la componente aleatoria del modelo [Canavos, 1998].

Si derivamos SSE en función de θ_0 y θ_1 , e igualando a cero, obtenemos los estimadores por mínimos cuadrados de la ordenada a la pendiente y al origen, respectivamente:

$$\theta_1 = \frac{SS_{xy}}{SS_{xx}} \quad (2.8)$$

$$\theta_0 = \bar{y} - \theta_1 \bar{x} \quad (2.9)$$

A la recta resultante $\hat{Y} = \hat{\theta}_0 + \hat{\theta}_1 X$, se la llama recta de regresión lineal de Y sobre X , es decir es el modelo de regresión lineal simple ajustado.

Siendo:

$$SS_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \quad (2.10)$$

$$SS_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 = n\sigma_x^2 \quad (2.11)$$

Un parámetro a estimar es la varianza de los errores σ^2 , su estimador se denomina varianza residual o CME y está dada por:

$$S^2 = \hat{S}_R^2 = \frac{SSE}{n-2} = \frac{\sum_{i=1}^n e_i^2}{n-2} = \frac{SS_{yy} - \beta_1 SS_{xy}}{n-2} \quad (2.12)$$

Es necesario tener una medida del ajuste de la recta de regresión y cuantificar el grado de asociación lineal entre las variables X y Y , al tener un mejor ajuste, se presentarán mejores pronóstico, como solución se utiliza el coeficiente de correlación lineal, dado por:

$$r = \frac{E[(X - E(X))(Y - E(Y))]}{\sqrt{VAR(X)VAR(Y)}} = \frac{\sqrt{VAR(X)}}{\sqrt{VAR(Y)}} \beta_1 \quad (2.13)$$

Este coeficiente puede tomar valores entre -1 y 1, siendo 0 el indicador de poca o ninguna relación entre variables, mientras más se acerque en valor absoluto a 1, mayor es su asociación lineal, un coeficiente positivo indica asociación lineal positiva, un coeficiente negativo indica variación en sentido opuesto.

Para estimar que tan bien se ajusta el modelo se utiliza el coeficiente de determinación lineal R^2 , que es la proporción de la variabilidad de la variable, expresada como:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SSE}{SS_{yy}} \quad (2.15)$$

Al ser aplicado este concepto a la regresión lineal, este coincide con el coeficiente de correlación lineal

$$R^2 = r^2 \quad (2.16)$$

El análisis de la regresión incluye los siguientes pasos:

- Planteamiento del problema
- Selección de variables potencialmente relevantes
- Recopilación de datos
- Especificación del modelo
- Elección del método de ajuste
- Ajuste del modelo
- Validación y crítica de modelos
- Utilización del modelo elegido para la solución del problema planteado

El planteamiento del problema es vital para una correcta selección de variables y para elegir el mejor método estadístico de análisis, la selección del conjunto de variables se realiza analizando que variables explican o predicen la variable respuesta, esta se denota por Y y la variable predictora está denotada por X_1, X_2, \dots, X_p , donde p denota el número de variables predictoras. La recopilación de datos consiste en n observaciones periódicas,

cada una de estas n observaciones constan de mediciones para cada una de las variables relevantes.

Tres métodos básicos para recopilar datos son los siguientes:

- Un estudio retrospectivo basado en datos históricos
- Un estudio observacional
- Un experimento diseñado

Un buen esquema de recopilación de datos puede garantizar un sistema simplificado y, en general, un modelo más aplicable [Montgomery et al., 2012].

La especificación del modelo que se considera que presenta la relación de la variable de respuesta con el conjunto de variables predictoras puede ser especificada inicialmente por la experticia de los profesionales en el área de estudio en función de su experiencia, sus juicios objetivos o subjetivos.

Una vez definido el modelo la siguiente tarea es estimar los parámetros del modelo en función de los datos recopilados, proceso conocido como el ajuste del modelo, el método más conocido es el de mínimos cuadrados (MC).

La selección de un modelo es un proceso iterativo hasta lograr un resultado satisfactorio, logrando un modelo que satisface las suposiciones de los datos y se ajusta razonablemente bien a ellos, este proceso iterativo se ilustra en la Figura 2.7.

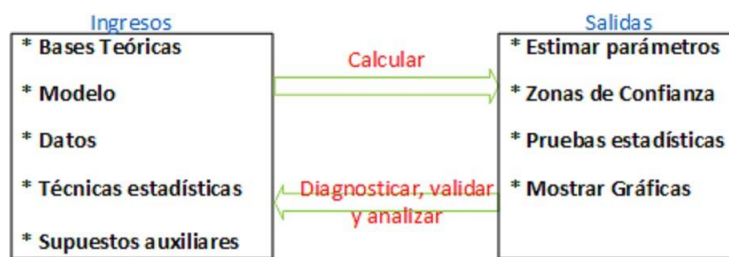


Figura 2.7: Ilustración esquemática de la naturaleza iterativa del proceso de regresión [Chatterjee and Hadi, 2012].

2.2.5. Redes neuronales

Las redes neuronales son sistemas formados por capas de nodos denominados neuronas, las cuales realizan algún tipo de acción u operación con los datos procedentes de la capa anterior para dar lugar a determinadas salidas. Estos sistemas tratan de reproducir el comportamiento de algún sistema real, siendo para ello entrenadas mediante un conjunto de datos conocidos. Estas redes han demostrado tener un buen comportamiento en el modelado de la carga y en la previsión de consumos. Para ello emplean diversos algoritmos de aprendizaje como pueden ser el algoritmo de retropropagación (BM) [Verdú, 2013].

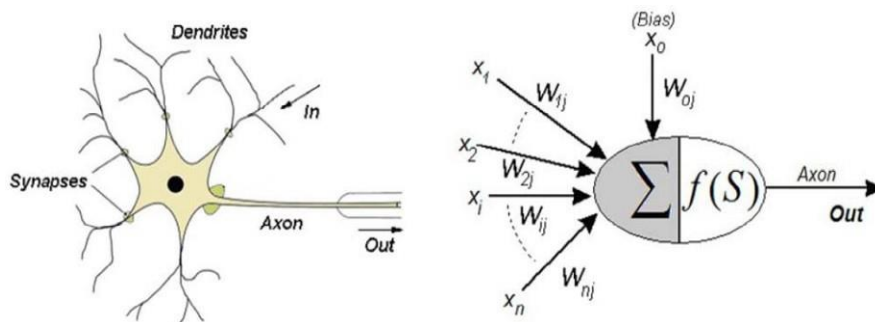


Figura 2.8: Neurona biológica genérica versus representación clásica de neurona artificial [Kattan et al., 2011]

2.2.6. Redes neuronales recurrentes (RNN's)

La característica fundamental de una RNN, es que la red contiene al menos una conexión de retroalimentación, por lo que las activaciones pueden ir en un bucle. Permite a las redes realizar un procesamiento temporal y aprender secuencias, por ejemplo, realizar un reconocimiento / reproducción de secuencias o una asociación / predicción temporal. Las arquitecturas RNN pueden tener muchas formas diferentes. Un tipo común consiste en un MLP (perceptrón multicapa) estándar, más bucles agregados. Estos pueden aprovechar las poderosas capacidades de mapeo no lineal del MLP y tener algún tipo de memoria.

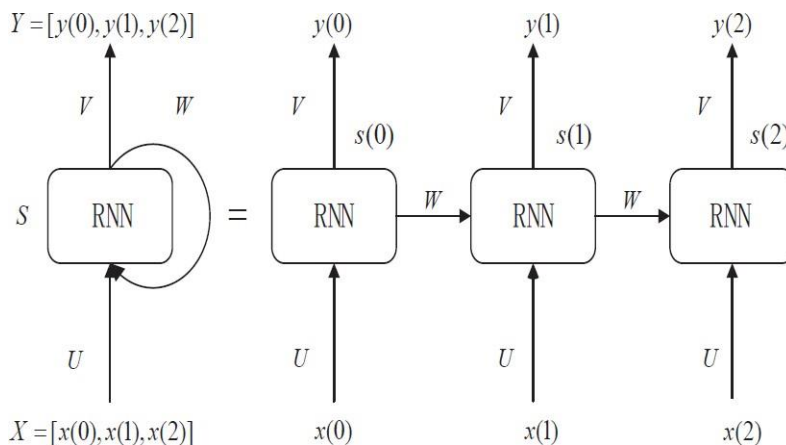


Figura 2.9: Estructura de una RNN [Ye et al., 2019]

En la Figura 2.9, los vectores de secuencia $X = [x(0), x(1), x(2)]$ pasan hacia la RNN uno por uno según el paso de tiempo establecido. Obviamente, esto es diferente de la tradicional red feedforward [Ye et al., 2019].

Para arquitecturas simples y funciones de activación deterministas, el aprendizaje se puede lograr utilizando procedimientos *GD* (gradiente descendiente) similares a los que conducen al algoritmo de retropropagación para redes de retroalimentación.

Las RNN están diseñadas para utilizar información secuencial de datos de entrada con conexiones cíclicas entre bloques de construcción como perceptrones, unidades de memoria a largo plazo a corto plazo (LSTM) o unidades recurrentes con compuerta (GRU). Los dos últimos se utilizan para eliminar los inconvenientes de los RNN regulares, como el problema de la desaparición / explosión del gradiente y la dependencia a largo y corto plazo [Fu et al., 2019].

2.2.7. Long short-term memory (LSTM)

Las redes LSTM son variantes de las RNN creadas para abordar el problema del desvanecimiento del gradiente. El problema del desvanecimiento del gradiente es causado por componentes de memoria que están demasiado lejos del paso actual y recibirían pesos más bajos debido a su distancia. Las LSTM son una variante de los RNN que contienen un componente de memoria, llamada puerta de olvido. Ese componente se puede utilizar para evaluar cómo los elementos antiguos y recientes afectan los pesos

y los sesgos, dependiendo de dónde se coloque la observación en una secuencia [Capelo, 2018].

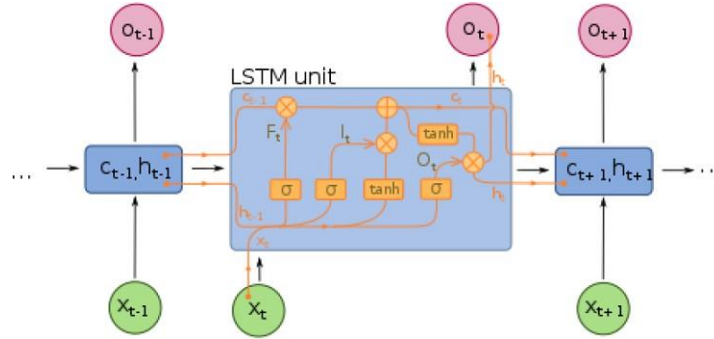


Figura 2.10: Estructura de LSTM [Mike Bernico, 2018]

Las LSTM son increíblemente efectivos en una variedad de problemas, como predecir el próximo valor en una serie de tiempo [Mike Bernico, 2018]. Comparados con otras redes neuronales, el modelo LSTM tiene tres unidades multiplicativas, es decir, Input gate, Output gate y Forget gate:

- **Input gate:** es la encargada de memorizar la información del presente, añadir nuevos elementos a la memoria.
- **Output gate:** es la encargada de abandonar información, permite crear el estado oculto actualizado.
- **Forget gate:** escoge que información olvidar del pasado, eliminar elementos de la memoria.

En la Figura 2.10, f_t es forget gate, i_t es input gate y O_t es output gate, que juntos controlan el estado celular, en tanto que C_t almacena información histórica y futura. x_t y h_t , respectivamente, representan la información de entrada y salida de la LSTM [Xiao, 2019].

El primer paso de la celda LSTM es identificar la información que no se requiere para omitirla. Este proceso de identificación y exclusión es decidido por la función sigmoideal, que toma la salida de la última unidad LSTM (h_{t-1}) en el tiempo $t-1$ y la entrada actual (x_t) en el tiempo t . Además, la función sigmoideal determina que parte de la salida anterior debe ser eliminada. Todo este proceso es llevado a cabo por forget gate (f_t), donde f_t es un vector de

valores entre 0 y 1, correspondiente a cada número en el estado de celda, (C_{t-1})

$$f_t = \sigma(\omega_f [h_{t-1}, x_t] + b_f) \quad (2.17)$$

Donde σ es la función sigmoideal, ω_f y b_f son las matrices de pesos y bias, respectivamente, del proceso forget gate.

El segundo paso es almacenar y decidir sobre la nueva entrada (x_t) entorno a la actualización del estado de la celda. Este proceso tiene dos partes, en primer lugar, se decide si la información debe actualizarse o ignorarse (0 o 1), encargado por la función sigmoideal, y, en segundo lugar, dependiendo de los valores que pudieron continuar, la capa tanh les otorga pesos a estos valores, según su nivel de importancia (-1 a 1). Los dos valores se multiplican para actualizar el estado de la celda. Esta nueva memoria se adiciona a la memoria anterior (C_{t-1}) dando como resultado C_t

$$i_t = \sigma(\omega_i [h_{t-1}, x_t] + b_i) \quad (2.18)$$

$$N_t = \tanh(\omega_n [h_{t-1}, x_t] + b_n) \quad (2.19)$$

$$C_t = C_{t-1}f_t + N_t i_t \quad (2.20)$$

donde C_{t-1} y C_t son las celdas de estado en el tiempo $t - 1$ y t , mientras que ω y b son matrices de pesos y bias, respectivamente de las celdas de estados.

Por último, los valores de salida (h_t) son basados por el estado de la celda de salida (o_t), pero siendo una versión filtrada. Pues, una capa sigmoideal decide que parte de la celda de estado llega a la salida, luego output gate se multiplica por los nuevos valores creados por la capa tanh desde el estado de la celda (C_t), que varía entre -1 y 1.

$$o_t = \sigma(\omega_o [h_{t-1}, x_t] + b_o) \quad (2.21)$$

$$h_t = o_t \tanh(C_t) \quad (2.22)$$

Donde ω_o y b_o corresponde a las matrices de pesos y bias, respectivamente, del output gate [Le et al., 2019].

Capítulo 3

Desarrollo del sistema

El capítulo describe los procedimientos seguidos para el desarrollo del sistema de pronóstico de fallas: la regresión lineal que presenta un baseline para el sistema, y el procedimiento de diseño de la red LSTM con las pruebas realizadas de optimización, balanceo de datos y redes neuronales creadas para los modelos de envolventes diaria, semanal y mensual.

3.1. Metodología del desarrollo

Este proyecto de desarrollo se basa principalmente en una metodología de investigación bibliográfica a través de libros, proyectos de investigación y artículos científicos, los cuales se fundamentan en la adquisición de señales de máquina en estado normal y estado de falla, para luego de una revisión de la calidad de la información se procede a una selección de parámetros, variables de interés y preparación de información para la aplicación de algoritmos de machine learning mediante aprendizaje supervisado y comenzar la primera etapa de pronóstico mediante regresión y así obtener un **baseline** para el proyecto. Como segunda etapa se aplicarán redes neuronales artificiales con LSTM para las etapas de entrenamiento y validación del pronóstico del sistema.

Como plataforma de desarrollo se ha escogido a Python al ser un lenguaje de programación de alto nivel, robusto y con librerías necesarias para proyectos de machine learning. Las Figuras (3.1 y 3.2), indican los procedimientos a seguir para las dos etapas de este proyecto.

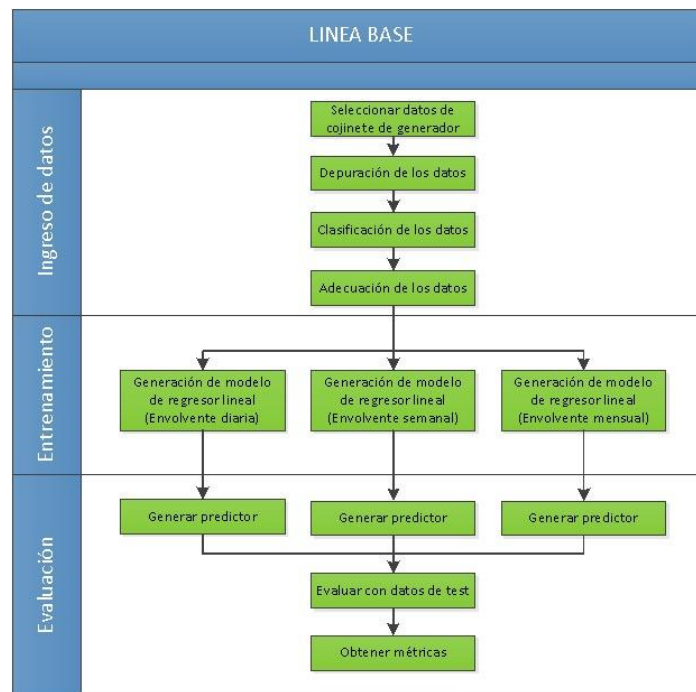


Figura 3.1: Metodología propuesta mediante regresores a desarrollarse

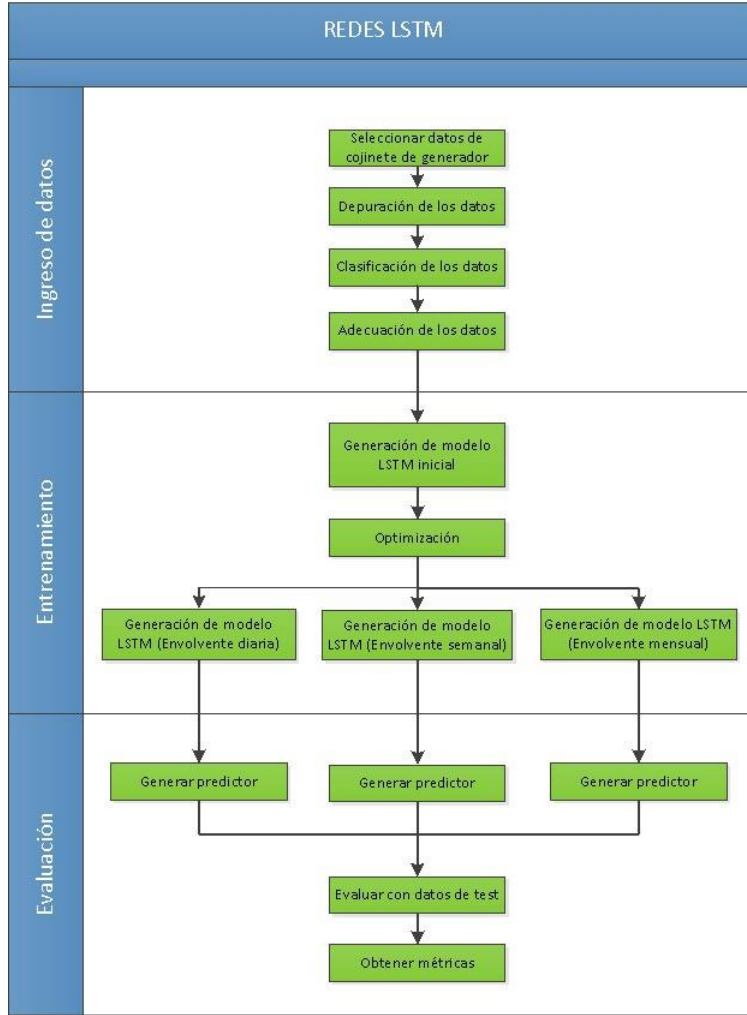


Figura 3.2: Metodología propuesta mediante RNN con LSTM a desarrollarse

3.2. Método de aprendizaje

Dentro de los métodos de aprendizaje de machine learning existen el supervisado y el no supervisado. El aprendizaje supervisado se utiliza para modelar la correspondencia entre un conjunto de N ejemplos o patrones de entrada $\{X^1, X^2, \dots, X^N\}$, y las salidas deseadas $\{t^1, t^2, \dots, t^N\}$, minimizando

para ello la función de error que mide la diferencia existente entre estas salidas y las obtenidas por la red $\{y^1, y^2, \dots, y^N\}$ [Méndez and J.T., 2008]. Es decir, se aprenden relaciones que asocian entradas con salidas, por lo que se ajustan a un conjunto de ejemplos de los que se conoce su relación entre la entrada y salida deseada. Los métodos no supervisados tratan de encontrar patrones o características significativas en los datos de entrada, en estos casos no se dispone de una salida para comparar el rendimiento del método por lo que la única información que tendrá el sistema es la que proporcionen los datos de entrada. En este trabajo se hace uso del método de aprendizaje supervisado.

3.3. Estructura y validación de la base de datos

El sistema tiene como base el uso de 54934 datos con registro horario, teniendo como fecha inicial el 01-01-2015 y final el 07-04-2021. Esta información se la tiene en un archivo .csv para el procesamiento, Figura 3.3.

Fecha	Año	Mes	Día	P ACTIVA	T AXIAL	T RADIAL	T GENERADOR	ESTADO
1/1/2015	2015	1	1	2.08	58	49	45.96	0
2/1/2015	2015	1	2	3.714	62	51.5	47.8	0
3/1/2015	2015	1	3	4.047	61	51.7	47.8	0
4/1/2015	2015	1	4	2.09	42	49	31.35	0
5/1/2015	2015	1	5	3.387	61	51.9	47.5	0
6/1/2015	2015	1	6	4.048	61	50.5	47.5	0
7/1/2015	2015	1	7	2.911	60	50.7	47.5	0
8/1/2015	2015	1	8	3.895	61	51.0	47.5	0

Figura 3.3: Parte del archivo principal csv.

Al tener una cantidad elevada de información se realizó un análisis visual de los datos cargados para poder apreciar en forma de histogramas como se encuentran distribuidos, la potencia activa de la unidad presenta una media de 3.83 y la desviación estándar es de 0.57, con lo que se establece que la mayor cantidad de datos de potencia activa (MW) se encuentran entre 3.26 y 4.4. Con respecto a la temperatura axial ($^{\circ}$), presenta una media de 65.36 y la desviación estándar es de 2.15, lo que establece que los rangos de temperatura se encuentran entre 63.21 y 67.51 grados, Figura 3.4.

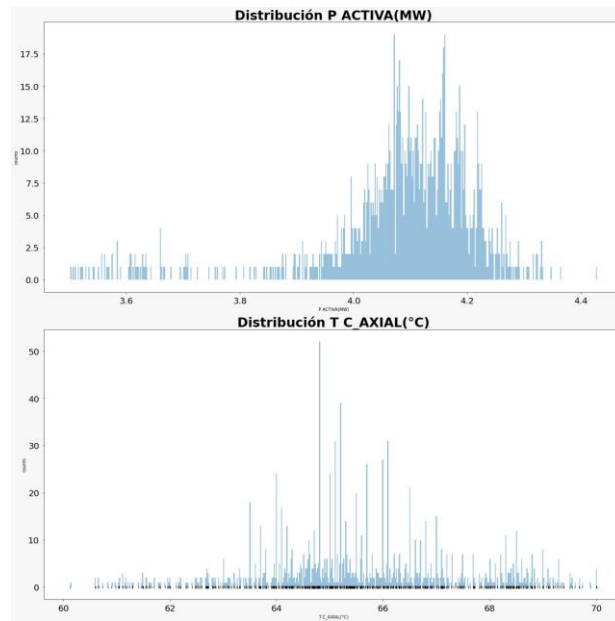


Figura 3.4: Histograma de distribución de los datos.

3.3.1. Limpieza de datos

La limpieza de datos también llamada lavado de datos (“data scrubbing”), trata de detectar y eliminar los errores e inconsistencias en los datos con el objetivo de mejorar su calidad. Los problemas de calidad de los datos pueden estar presentes en colecciones de datos tales como ficheros y bases de datos, producto de errores ortográficos en la entrada, información ausente y otros datos no válidos [López Porrero, 2009]. La limpieza de datos implicó la búsqueda, detección y corrección de valores duplicados, datos inconsistentes que se han grabado debido algún incidente del sensor al tomar la muestra y valores que se encuentran por fuera del estudio, esto con el objetivo de mejorar su calidad para el ingreso al sistema.

3.3.2. Selección de datos y parámetros

Para la selección de parámetros se analizó la variabilidad y tipo de correlaciones existentes entre las variables disponibles de potencia activa, temperatura de generador, temperatura de cojinete radial y temperatura de cojinete axial, seleccionando esta última como la variable de ingreso

al sistema por presentar la mejor calidad de información de falla del cojinete, Figura 3.5, con la característica de que un valor superior a los 68.5 grados implica una falla. Se cuenta con 45037 datos en estado 0 (valores de temperatura en estado normal) y 9897 datos en estado 1 (valores de temperatura en estado de falla).

Debido a la gran cantidad de información de temperatura axial del cojinete, se utilizó los picos que estos producen, formándose una envolvente de datos, en las que se generaron modelos diarios, semanales y mensuales de pronóstico tanto para regresión lineal como para la red LSTM.

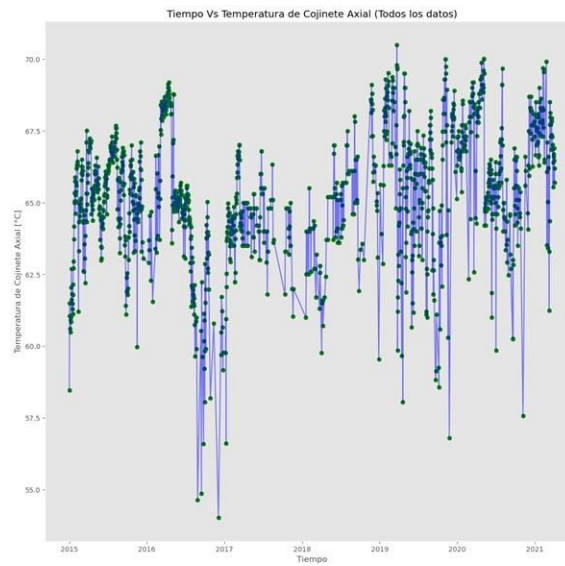


Figura 3.5: Medición histórica de temperatura axial 2015 a 2021.

Para apreciar la información de relevancia se visualizan los datos en estado normal (color azul) y en estado de falla (rojo), Figura 3.6, además de los datos únicamente en estado normal, Figura 3.7, y finalmente sólo los datos en falla en la Figura 3.8.

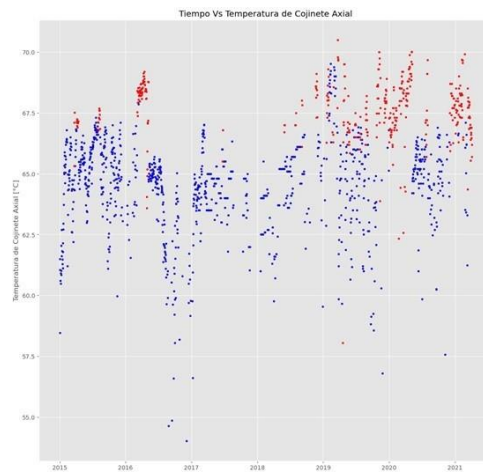


Figura 3.6: Mediciones históricas de temperatura en estado normal (azul) y en falla (rojo) 2015 a 2021.

En la Figura 3.6 se aprecian datos de falla en valores menores de 68.5 grados, esto es debido a que cuando la unidad se vió forzada a parar por un fallo repentino, la temperatura de la misma baja gradualmente y al encontrarse en la zona de falla, el sistema reconoce estas temperaturas que son inferiores al umbral establecido como valores de falla.

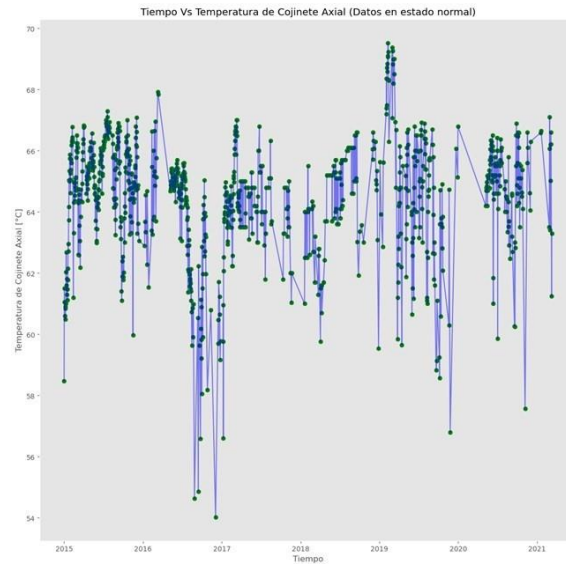


Figura 3.7: Mediciones históricas de temperatura en estado normal 2015 a 2021.

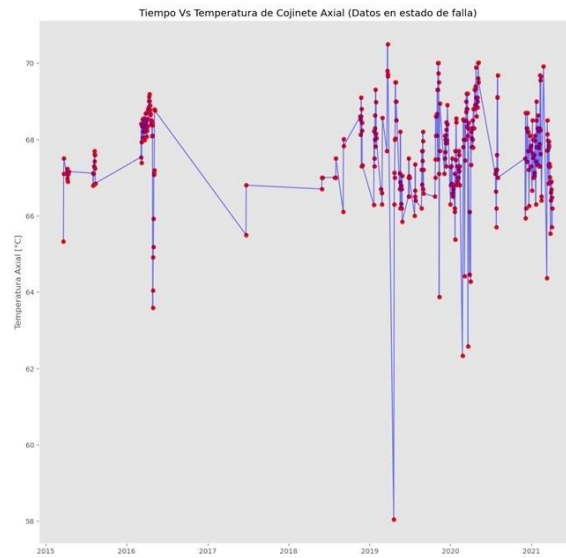


Figura 3.8: Mediciones históricas de temperatura en estado de falla 2015 a 2021.

3.4. DISEÑO DEL SISTEMA DE PRONÓSTICO MEDIANTE REGRESIÓN 40

3.3.3. Normalización de datos

La normalización es un proceso en el que se establecen una regla o conjunto de reglas para la conversión de estructuras de datos complejos en datos más simples, logrando de esta manera facilitar el manejo y operación de los mismos debido a que algunas estructuras de datos presentan diferentes escalas numéricas como también diferentes distribuciones [Aldaz et al., 2009, Gil Gonzalez et al., 2014].

Para la normalización de datos se utiliza el valor máximo y mínimo de la base de datos para obtener los valores entre 0 y 1 [Gil Gonzalez et al., 2014, Zhang and You, 2015].

$$S_{ij}^{\prime} = \frac{S_{ij} - \min_j}{\max_j - \min_j} \quad (3.1)$$

Donde S_{ij}^{\prime} es el dato transformado del conjunto de datos, S_{ij} es el dato del conjunto de datos, \min_j es el valor mínimo del conjunto de datos, \max_j es el valor máximo del conjunto de datos.

3.4. Diseño del sistema de pronóstico mediante regresión

Se considera el uso de las series temporales, las cuales son una lista de estados temporalmente ordenados uno detrás de otro. La utilidad de las series del tiempo está en su análisis para realizar pronósticos, fundamentado en una secuencia de observaciones que son medidas en determinados tiempos ordenados cronológicamente, logrando una dependencia entre los datos [Ma, 2007].

Se divide los datos en tres segmentos: un 80 % para entrenamiento, 10 % para validación y el restante 10 % únicamente para evaluación. Estos mismos segmentos serán utilizados para el sistema de pronóstico mediante LSTM.

Con los datos de temperatura axial del cojinete se ajustan las rectas de regresión lineal para los años que abarca este estudio, se han aplicado las técnicas de regresión para obtener 3 modelos:

- Modelo de regresión lineal con datos diarios
- Modelo de regresión lineal con datos semanales
- Modelo de regresión lineal con datos mensuales

3.4. DISEÑO DEL SISTEMA DE PRONÓSTICO MEDIANTE REGRESIÓN

El objetivo de este ajuste es predecir una fecha en la que se presentará una falla de cojinete en función de los valores de temperatura conocidos (variable independiente) y la temperatura predicha (variable dependiente).

En general los pasos que se han seguido para el desarrollo de los algoritmos de regresión han sido los siguientes:

- **Carga y filtrado de datos:** Se realiza la importación de librerías y la lectura de los archivos .csv (que contiene la información de la BD), en cuanto a la filtración de datos se estableció trabajar con todos aquellos valores que superan el umbral de 1 MW de potencia, debido a que las fallas en el cojinete se producen cuando la unidad de generación está operativa con potencias entre los 3 MW a 4 MW, (conocimiento experto).
- **Análisis visual:** ha permitido apreciar de forma gráfica la distribución de las variables, características y su comportamiento en el tiempo, para de esta manera delimitar los rangos de datos del proyecto tanto para regresión como con la red LSTM, con el uso de la media y desviación estándar de las mismas.
- **Aplicación del regresor:** se realizó el escalamiento de todos los segmentos: entrenamiento, validación y prueba, se creó el objeto de regresión lineal y a estos se aplican ventanas para abarcar todo el conjunto de datos, en que las individualmente se realiza la regresión lineal, obteniendo una recta por cada ventana hasta el final de los datos, y estas se ordenan por su fecha para obtener el modelo equivalente del sistema generado mediante regresión. Una vez realizado este procedimiento se suman y promedian las métricas de evaluación: *RMSE* y el R^2 de cada una de las ventanas individuales para obtener los valores de métricas resultantes del modelo. Con la última ventana se obtiene los coeficientes de la ecuación lineal y a partir de ellos se realiza el pronóstico.

3.4.1. Modelo de regresión lineal con datos diarios

Una vez realizado el procedimiento explicado en 3.4, se presenta en la Figura 3.9 el modelo de regresión lineal con datos diarios de 2015 a 2021, las métricas del modelo son iguales al promedio del *RMSE* y el R^2 de cada una de las rectas de regresión que se generan por cada ventana.

3.4. DISEÑO DEL SISTEMA DE PRONÓSTICO MEDIANTE REGRESIÓN

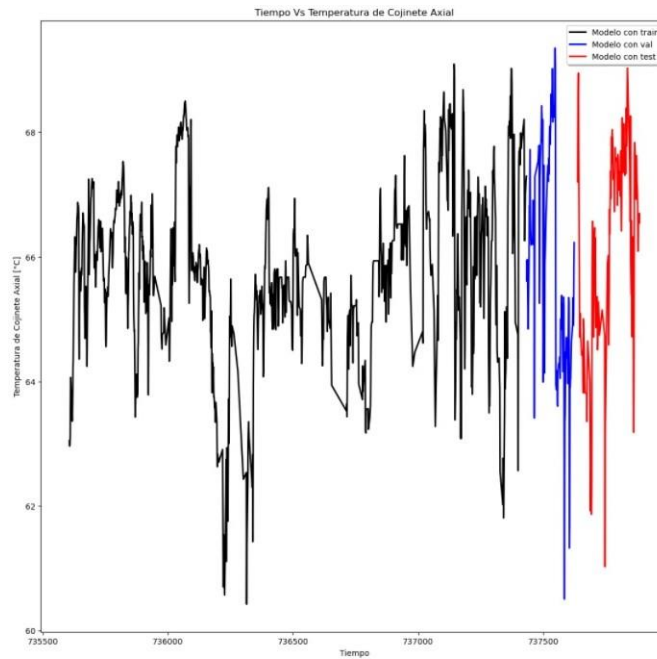


Figura 3.9: Modelo de regresión lineal con datos diarios

Los valores resultantes son:

- $RMSE$: 0.01
- R^2 : -1.51
- Pendiente: 0.006495591569988336
- Término independiente: -4726.390525704896

La ecuación del modelo de la última ventana con los datos diarios es:
 $y = (0,006495591569988336)x + (-4726,390525704896)$

La temperatura de fallo es de $68.5\text{ }^{\circ}\text{C}$ por lo que para poder pronosticar la fecha de un posible fallo, reemplazamos este valor en la ecuación del modelo lineal de datos diarios, el mismo que indica un fallo para el: **2022-01-21**.

3.4.2. Modelo de regresión lineal con datos semanales

Una vez realizado el procedimiento explicado en 3.4, se presenta en la Figura 3.10 el modelo de regresión lineal con datos semanales de 2015 a

3.4. DISEÑO DEL SISTEMA DE PRONÓSTICO MEDIANTE REGRESIÓN

2021, las métricas del modelo son iguales al promedio del $RMSE$ y el R^2 de cada una de las rectas de regresión que se generan por cada ventana.

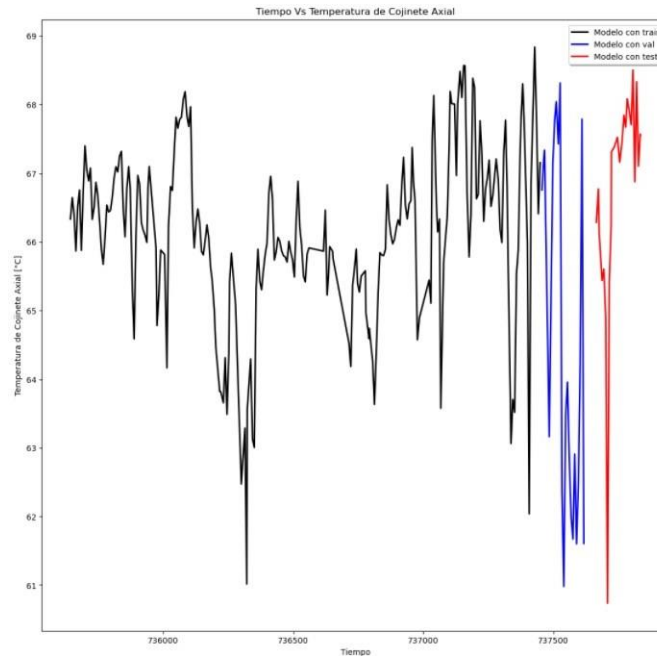


Figura 3.10: Modelo de regresión lineal con datos semanales

Los valores resultantes son:

- $RMSE$: 0.03
- R^2 : -2.73
- Pendiente: 0.009251500075567737
- Término independiente: -6758.917046350458

La ecuación del modelo de la última ventana con los datos semanales es:
 $y = (0,009251500075567737)x + (-6758,917046350458)$

La temperatura de fallo es de $68.5\text{ }^{\circ}\text{C}$ por lo que para poder pronosticar la fecha de un posible fallo, reemplazamos este valor en la ecuación del modelo lineal de datos semanales, el mismo que indica un fallo para el: **2021-07-08**.

3.4. DISEÑO DEL SISTEMA DE PRONÓSTICO MEDIANTE REGRESIÓN

3.4.3. Modelo de regresión lineal con datos mensuales

Una vez realizado el procedimiento explicado en 3.4, se presenta en la Figura 3.11 el modelo de regresión lineal con datos mensuales de 2015 a 2021, las métricas del modelo son iguales al promedio del $RMSE$ y el R^2 de cada una de las rectas de regresión que se generan por cada ventana.

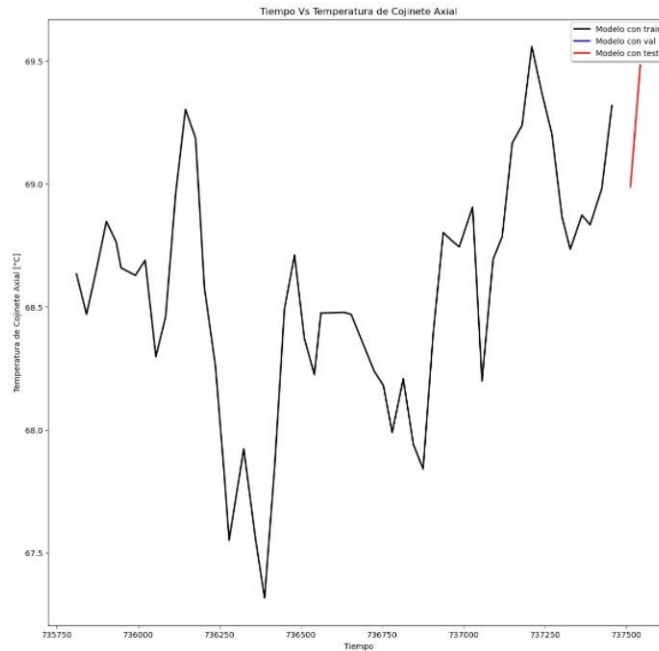


Figura 3.11: Modelo de regresión lineal con datos mensuales

Los valores resultantes son:

- $RMSE$: 0.09
- R^2 : -0.91
- Pendiente: 0.13483728616251506
- Terminio independiente: -99424.74904918502

La ecuación del modelo de la última ventana con los datos mensuales es:

$$y = (0,13483728616251506)x + (-99424,74904918502)$$

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM45

La temperatura de fallo es de $68.5\text{ }^{\circ}\text{C}$ por lo que para poder pronosticar la fecha de un posible fallo, reemplazamos este valor en la ecuación del modelo lineal de datos mensuales, el mismo que indica un fallo para el: **2021-03-27**.

3.5. Diseño de sistema de pronóstico mediante RNN con LSTM

Como segunda parte de este proyecto se presenta un modelo basado en una red LSTM para el pronóstico de falla del cojinete. La principal ventaja con respecto a las tradicionales RNN son que las redes LSTM pueden mantener en la memoria las dependencias de tiempo entre las entradas para un período más largo. Las redes LSTMs son sensibles a la escala de los datos de entrada, por lo que se ha aplicado la normalización de datos descrita en 3.3.3.

El modelo inicial de la red LSTM tiene 5 datos de entrada, una capa con 4 bloques LSTM o neuronas, una capa oculta con una densidad de 2 neuronas y una capa de salida que hace un solo pronóstico de valor. La primera red se creó para 100 épocas o iteraciones “**epochs**”, un tamaño de lote por muestras “**batch**” de 12 y el optimizador Adagrad (Figura 3.12).

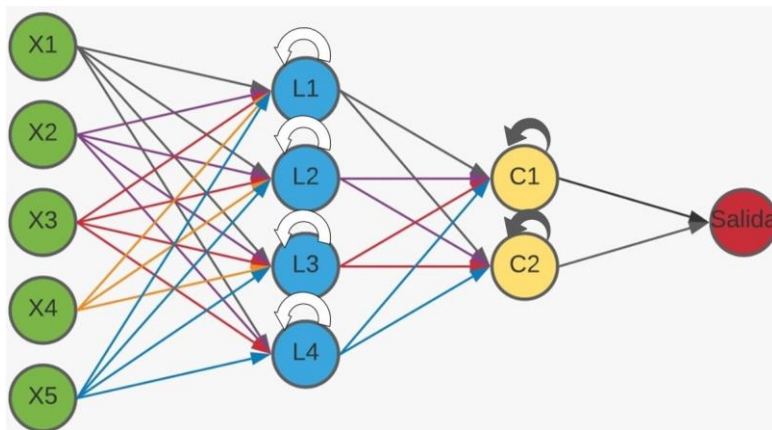


Figura 3.12: Modelo inicial de la red neuronal con LSTM

La Figura 3.13 muestra la función de pérdida del primer modelo sin optimizar la red neuronal.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM46

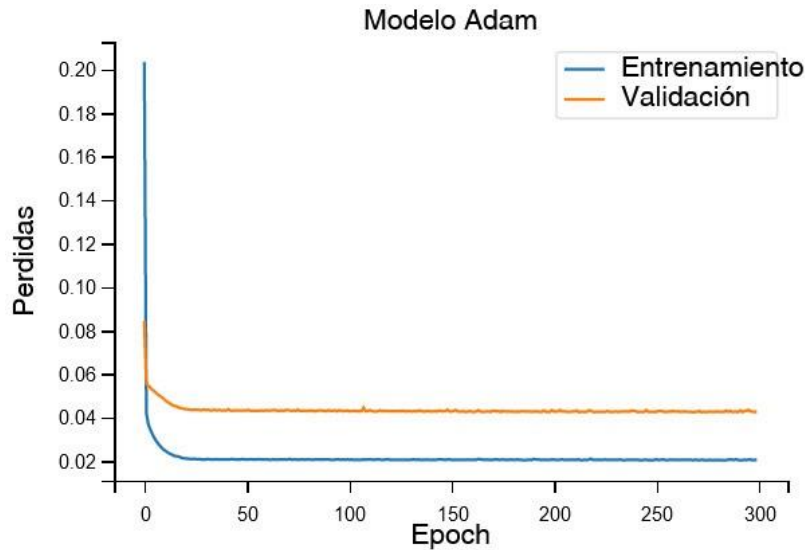


Figura 3.13: Función de pérdida del modelo inicial

Para el desarrollo de los modelos de pronóstico de falla se mantienen los mismos criterios que se han planteado para la regresión lineal: la división en tres segmentos, un 80 % para entrenamiento, 10 % para validación y el restante 10 % únicamente para evaluación. Así como se conserva el criterio de pronosticar una fecha en la que se presentará una falla de cojinete en función de los valores de temperatura conocidos (variable independiente) y la temperatura predicha (variable dependiente).

En general los pasos que se han seguido para el desarrollo de los modelos LSTM han sido los siguientes:

- **Carga y filtrado de datos:** Se realiza la importación de librerías y la lectura de los archivos .csv, (que contiene la información de la BD), en cuanto a la filtración de datos se estableció trabajar con todos aquellos valores que superan el umbral de 1 MW de potencia, debido a que las fallas en el cojinete se producen cuando la unidad de generación está operativa con potencias entre los 3 MW a 4 MW, (conocimiento experto).
- **Aplicación de balanceo de datos:** Se han aplicado las técnicas de sobremuestreo para aumentar los datos de la clase minoritaria

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM47

e igualar a la de la mayoritaria, submuestreo para reducir la clase mayoritaria a la misma cantidad de datos de la clase minoritaria y combinación de ambas técnicas para obtener un modelo con ambas técnicas de balanceo. Estas técnicas han sido aplicadas a los segmentos de entrenamiento y validación, mas no al de evaluación el cual es el mismo utilizado para los modelos de regresión como para los LSTM.

- **Diseño de red LSTM:** Se tomó el 80 % de los datos para el entrenamiento, y se crean dos matrices $[X, Y]$ en la que X lleva datos de ingreso y Y llevará los datos que se pronosticarán, aplicando ventanas para abarcar todo el conjunto de datos. A estos datos se le realizan un escalamiento para poder trabajar con datos normalizados entre 0 y 1, luego se obtienen los datos de entrenamiento y prueba. Se cambia el formato de la estructura de los datos para que estos puedan ser recibidos por la red LSTM. Se crea una red con las características de la red optimizada de la Tabla 3.8 y se entrena el modelo con el que se obtiene la curva de función de pérdidas. Una vez realizado este procedimiento se obtienen las métricas de evaluación: *RMSE* y el R^2 resultantes del modelo. Con la ultima ventana de datos se aplica el algoritmo free-running para obtener el pronóstico.

3.5.1. Optimización de la red neuronal

Para la optimización de la red Neuronal se entrenan los modelos ajustando los pesos adecuados para minimizar el error de la salida del modelo en comparación con la salida esperada. Esto lleva el nombre de algoritmo de optimización de descenso de gradiente estocástico. Mientras que el algoritmo de retropropagación ajusta los pesos del modelo usando una regla específica del cálculo que asigna el error proporcionalmente a cada peso de la red [Aggarwal, 2018].

Se definen 3 etapas para la optimización de la red neuronal.

Elección de algoritmo optimizador

Se han empleado 3 algoritmos de optimización:

Adagrad (Adaptative Gradient Algorithm), el cual utiliza diferentes tasas de aprendizaje para las variables teniendo en cuenta el gradiente acumulado en cada una de ellas [Duchi et al., 2011].

RMSprop (Root Mean Square Propagation) que es una variación de Adagrad en la que, en lugar de mantener un acumulado de los gradientes, se utiliza el concepto de "ventana" para considerar solo los gradientes más

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM48

recientes.

Adam (Adaptative Moment Optimization) combina la metodología de Momentum y RMSProp, calculando una combinación lineal entre el gradiente y el incremento anterior, y considera los gradientes recientemente aparecidos en las actualizaciones para mantener diferentes tasas de aprendizaje por variable [Zaheer and Shaziya, 2020].

Para la elección del optimizador se realizaron pruebas de combinación de batch de 5, 20, 30 y 40 con combinaciones de 50, 100, 200 y 300 épocas para los tres optimizadores, con resultados que se presentan en las Tablas 3.1, 3.2 y 3.3.

Tabla 3.1: Resultados batch vs épocas con Adam

		EPOCAS			
		50	100	200	300
BATCH	5	0,22855002	0,22978504	0,2310941	0,23184110
	20	0,22870552	0,23214159	0,22804079	0,2334342
	30	0,23413663	0,23030372	0,23237898	0,23237898
	40	0,23134223	0,23506035	0,22921957	0,23068565

Tabla 3.2: Resultados batch vs épocas con RMSPROP

		EPOCAS			
		50	100	200	300
BATCH	5	0,23217636	0,23656548	0,23309042	0,23875966
	20	0,22914334	0,23090549	0,23042165	0,22984728
	30	0,23076820	0,23025648	0,22944804	0,22944804
	40	0,23278686	0,23111269	0,22973334	0,22933000

Tabla 3.3: Resultados batch vs épocas con Adagrad

		EPOCAS			
		50	100	200	300
BATCH	5	0,24589160	0,24619475	0,23984805	0,25285870
	20	0,38635317	0,34391381	0,25040710	0,25092334
	30	0,52076042	0,33728298	0,24491529	0,27541571
	40	0,40874595	0,30788379	0,24581004	0,25397542

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM49

Podemos observar que el mejor resultado se ha presentado con el optimizador Adam, con un batch de 40 y 300 épocas, la cual su función de pérdida se presenta en la Figura 3.14.

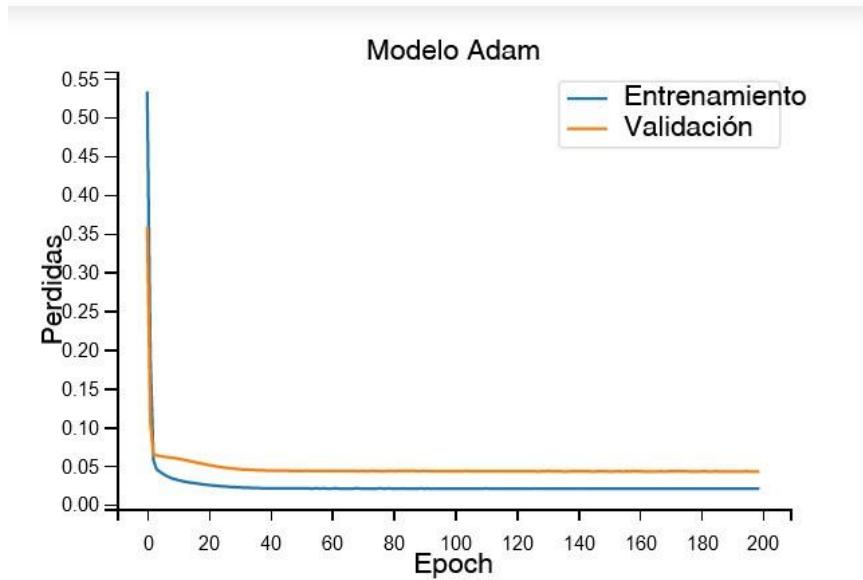


Figura 3.14: Función de pérdida con optimizador Adam

Optimización de capas neuronales

Para este punto se buscó la mejor combinación de neuronas de capa Dense (2, 8, 16, 32, 64 y 256) vs la cantidad de neuronas LSTM en el rango de (4, 8, 16, 64, 128 y 256), que presenten el menor error RMSE, como se puede apreciar en la Tabla 3.4 y el más alto coeficiente R^2 , Tabla 3.5.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM50

Tabla 3.4: Resultados Neuronas capa Dense vs Neuronas LSTM para RMSE

		ERROR RMSE					
		Neuronas LSTM					
		4	8	16	64	128	256
Neuronas capa Dense	2	0,2324291	0,2313868	0,2298889	0,2288129	0,2295993	0,2302471
	8	0,2307268	0,2287414	0,2330597	0,2312280	0,2298334	0,2305060
	16	0,2274615	0,2298977	0,2306419	0,2325522	0,2302988	0,2307133
	32	0,2287231	0,2298852	0,2315180	0,2325695	0,2314069	0,2323396
	64	0,2307573	0,2286425	0,2292635	0,2360347	0,2310605	0,2314289
	256	0,2292839	0,2303927	0,2296038	0,2328983	0,2315514	0,2331785

Se puede notar que el menor error RMSE se produce con 16 neuronas capa Dense y 4 neuronas LSTM.

Tabla 3.5: Resultados Neuronas capa Dense vs Neuronas LSTM para coeficiente R^2

		COEFICIENTE R2					
		Neuronas LSTM					
		4	8	16	64	128	256
Neuronas capa Dense	2	0,366921915	0,372586789	0,380683939	0,386467904	0,382243096	0,378752135
	8	0,376161162	0,386851215	0,363481669	0,37344765	0,380982703	0,377354207
	16	0,393693618	0,380636505	0,376619947	0,366251102	0,378473317	0,376233736
	32	0,38694903	0,380703536	0,371875337	0,366156568	0,372477991	0,36740918
	64	0,375996057	0,387381252	0,384048616	0,347127641	0,37435516	0,37235853
	256	0,383939087	0,37796611	0,382218809	0,364362899	0,371693816	0,362832897

La mejor combinación para lograr el mayor coeficiente R^2 son 16 neuronas capa dense y 4 neuronas LSTM, mientras más cercano este coeficiente sea a 1, implica un mejor modelo, como comprobación de los mejores parámetros estimados se aplicó la técnica Grid Search, la cual es un método de ajuste de parámetros en aprendizaje supervisado que realiza todas las combinaciones posibles y nos indica la mejor opción, visualizándose mediante un mapa de calor (heatmap), como se puede apreciar en la Figura 3.15.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM51

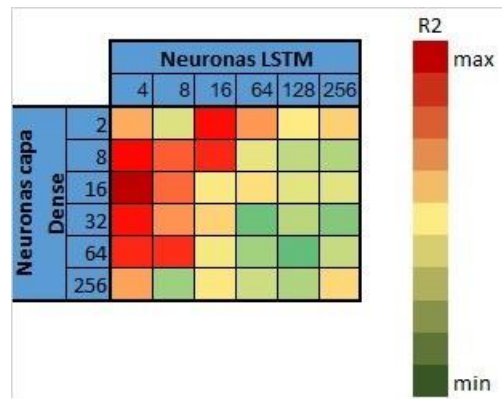


Figura 3.15: Visualización mediante mapa de calor de los mejores parámetros

Una vez realizadas las etapas de optimización de capas neuronales (capa dense, LSTM) y la elección del optimizador Adam (mejor combinación de batch y épocas), se puede apreciar la mejora de la función de pérdida con estas dos etapas de optimización aplicadas en la Figura 3.16.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM52

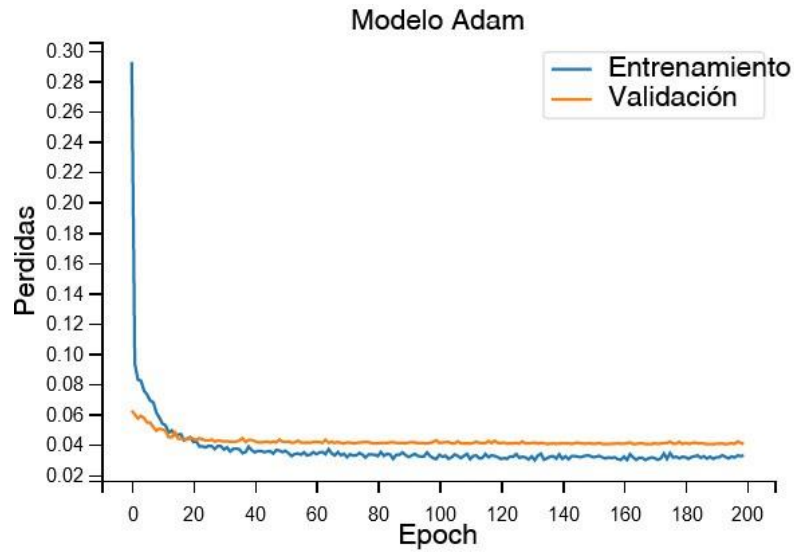


Figura 3.16: Función de pérdida con optimización de capas neuronales y algoritmo Adam

Optimización con Dropout

El dropout es una técnica de regularización que se basa en la eliminación de neuronas en las capas de la red neuronal que es aplicada en base a la probabilidad dada por la distribución de Bernoulli [Castañares, 2020]. Para este punto se buscó la mejor combinación dropout de neuronas de capa Dense (0, 0.1, 0.3, y 0.4) vs dropout en capas LSTM en el rango de (0, 0.1, 0.3, y 0.4) para buscar el menor error RMSE, como se puede apreciar en la Tabla 3.6, y el más alto coeficiente R^2 , Tabla 3.7.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM53

Tabla 3.6: Resultados aplicación dropout capa Dense vs capa LSTM para error RMSE

		ERROR RMSE				
		Dropout en capa LSTM				
		0	0,1	0,2	0,3	0,4
Dropout Dense	0	0,228498172	0,226192061	0,226197614	0,227051357	0,225670279
	0,1	0,22973487	0,226143171	0,224899769	0,22592665	0,225826058
	0,2	0,229902369	0,22898791	0,22797542	0,227070516	0,224780061
	0,3	0,230854655	0,22715195	0,226852563	0,225650218	0,225860853
	0,4	0,227362798	0,225707533	0,2248252	0,225878605	0,226492455

Se puede apreciar en la Tabla 3.6 que el mínimo error RMSE se produce al aplicar dropout en la capa dense de 0.2 y en la capa LSTM de 0.4.

Tabla 3.7: Resultados aplicación dropout capa Dense vs capa LSTM para coeficiente R^2

		COEFICIENTE R2				
		Dropout en capa LSTM				
		0	0,1	0,2	0,3	0,4
Dropout Dense	0	0,388154305	0,40044205	0,400412614	0,395877998	0,403204992
	0,1	0,381513411	0,400701207	0,40727333	0,401848255	0,402380781
	0,2	0,380611208	0,38552877	0,390950633	0,395776043	0,40790415
	0,3	0,375469401	0,395342578	0,396935409	0,403311095	0,402196606
	0,4	0,394219543	0,403007939	0,40766632	0,402102631	0,398848511

La mejor combinación se nota en la Tabla 3.7, que el mayor coeficiente R^2 se produce al aplicar dropout con 0.2 en capa dense y 0.4 en la capa LSTM.

Un cuadro de resumen entre la red neuronal inicial y la red neuronal optimizada se presenta en la Tabla 3.8.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM54

Tabla 3.8: Parámetros de red neuronal inicial y optimizada

	Red Neuronal Inicial	Red Neuronal Optimizada
Neuronas LSTM	4	4
Neuronas Dense	2	16
Dropout LSTM	0	0,4
Dropout Dense	0	0,2
Optimizador	Adam	Adam
Epocas	100	200
Batch	12	20
RMSE	0,230	0,224
R^2	0,379	0,407

3.5.2. Balanceado de clases normales y de falla

La unidad de generación opera en gran parte del año en condiciones normales y con algunos incidentes de falla por lo que la cantidad de datos que se dispone en estado normal de trabajo difieren de los estados en falla, (Figura 3.17). El contar con una base de datos no balanceada para un sistema del tipo supervisado, implica que estas clases no se encuentran distribuidas equitativamente, lo que produce efectos no deseados en el proceso de pronóstico, afectando a la capacidad de generalización de los algoritmos y perjudicando a las clases minoritarias, por lo que se aplicaron técnicas de balanceo de datos con sobremuestreo, submuestreo y una combinación de estas para obtener el mejor modelo posible.

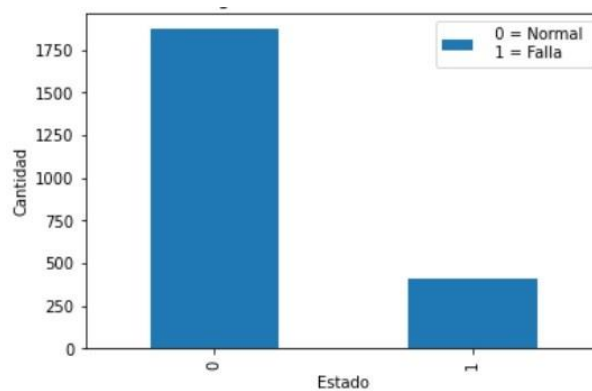


Figura 3.17: Histograma de estados del generador

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM55

Para el sobremuestreo (oversampling) se aplicó **SMOTE** (synthetic minority oversampling technique), la cual con algoritmos de interpolación incrementa en forma sintética los ejemplos de la clase minoritaria, teniendo de esta manera ambas clases igualadas. Para el submuestreo se aplicó el algoritmo **Cluster Centroids** el cual toma muestras de la clase mayoritaria mediante la sustitución de un grupo de muestras mayoritarias por el centroide, logrando representar el grupo de datos con un número reducido de muestras.

3.5.3. Modelado LSTM con envolvente diaria

Para el caso de la envolvente diaria se obtuvieron 1871 datos en estado 0 (estado normal) y 413 datos en estado 1 (falla), a los cuales se aplicaron las técnicas de balanceo de datos.

Diseño de red neuronal con sobremuestreo

Una vez realizado el procedimiento explicado en 3.5, se presenta en la Figura 3.18 la curva de función de pérdidas para este modelo, donde se ha aplicado el sobremuestreo con la técnica de SMOTE, con un resultado del estado normal (0) de 1871 a 1871 datos y de estado de falla (1) de 413 a 1871, quedando con los datos balanceados.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM56

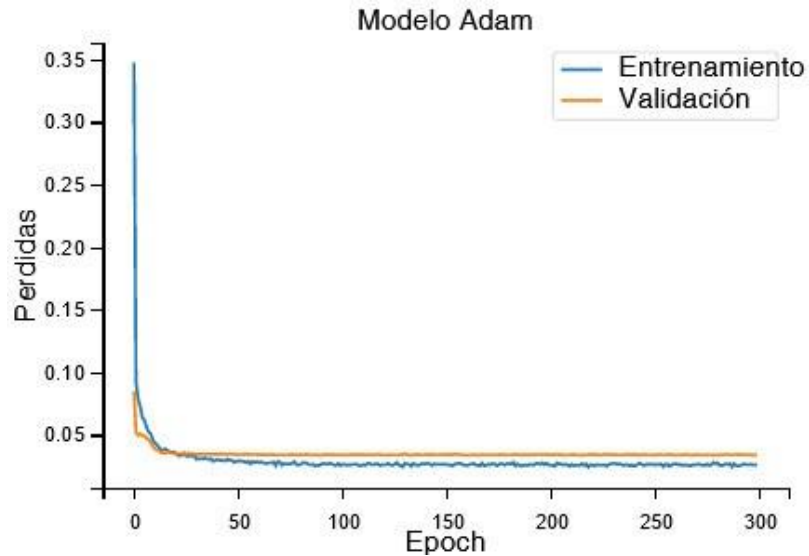


Figura 3.18: Función de pérdidas para envolvente diaria con sobremuestreo

Las métricas de evaluación de este modelo presentan un error $RMSE$: 0.1254 y un coeficiente R^2 : 0.5011.

Diseño de red neuronal con submuestreo

Para el desarrollo de este modelo se ha realizado el procedimiento explicado en 3.5, se presenta en la Figura 3.19 la curva de función de pérdidas para este modelo, donde se ha aplicado el submuestreo con la técnica de Cluster Centroids, con un resultado del estado normal (0) de 1871 a 413 datos y de estado de falla (1) de 413 a 413, quedando con los datos balanceados.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM57

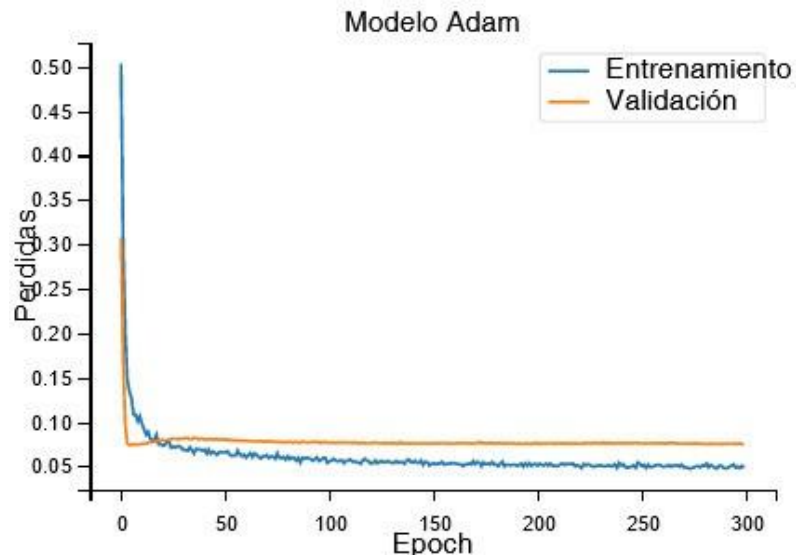


Figura 3.19: Función de pérdidas para envoltorio diaria con submuestreo

Las métricas de evaluación de este modelo presentan un error $RMSE$: 0.1752 y un coeficiente R^2 : 0.1250.

Diseño de red neuronal con sobre y submuestreo

Para el desarrollo de este modelo que combina las técnicas de sobre y sub muestreo, se aplicó el procedimiento explicado en 3.5, donde se realizó en una primera instancia el sobremuestreo con una relación del 60 % entre datos en estado normal y falla, es decir, sin llegar a tener los datos totalmente balanceados, con un resultado parcial del estado normal (0) de 1871 a 1305 datos y de estado de falla (1) de 413 a 783 datos, posteriormente se aplicó el submuestreo, con un resultado del estado normal (0) de 1305 a 783 datos y de estado de falla (1) de 783 a 783 datos, quedando con los datos balanceados, la Figura 3.20 presenta la curva de función de pérdidas para este modelo.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM58

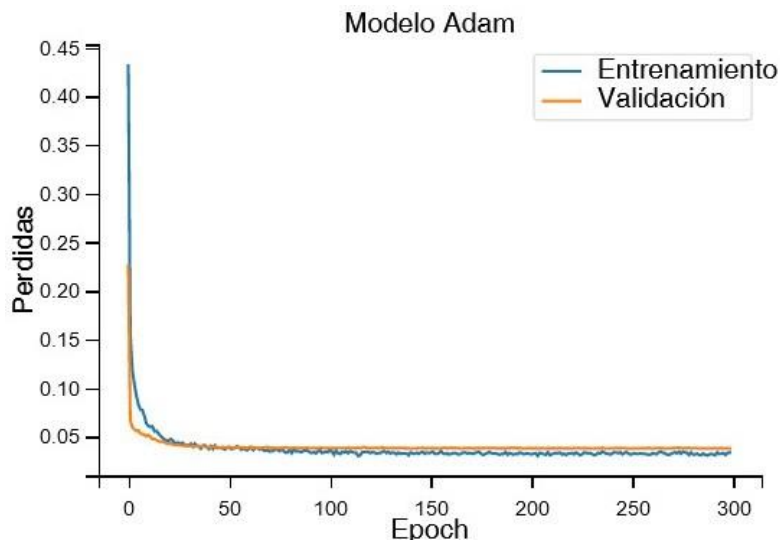


Figura 3.20: Función de pérdidas para envoltorio diaria con técnicas combinadas

Las métricas de evaluación de este modelo presentan un error *RMSE*: 0.2109 y un coeficiente R^2 : 0.3163.

3.5.4. Modelado LSTM con envoltorio semanal

Para el caso de la envoltorio semanal se obtuvieron 239 datos en estado 0 (estado normal) y 87 datos en estado 1 (falla), a los cuales se aplicaron las técnicas de balanceo de datos.

Diseño de red neuronal con sobremuestreo

Una vez realizado el procedimiento explicado en 3.5, se presenta en la Figura 3.21 la curva de función de pérdidas para este modelo, donde se ha aplicado el sobremuestreo con la técnica de SMOTE, con un resultado del estado normal (0) de 239 a 239 datos y de estado de falla (1) de 87 a 239, quedando con los datos balanceados.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM59

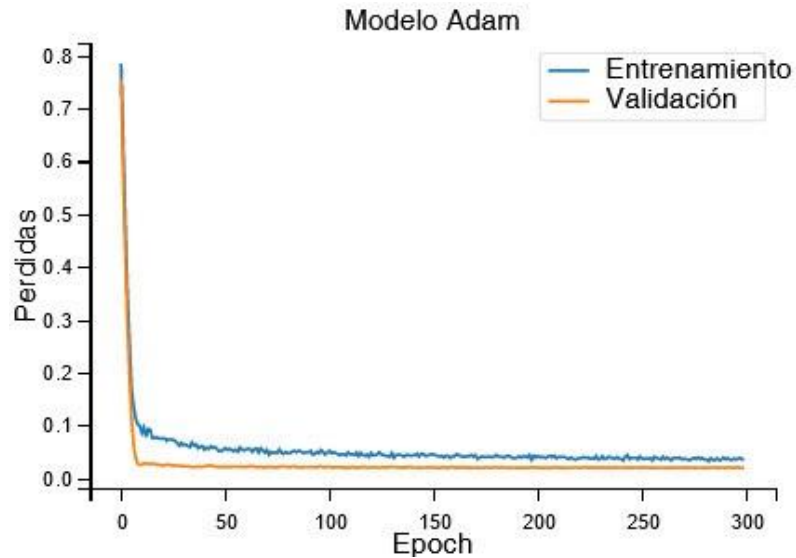


Figura 3.21: Función de pérdidas para envoltente semanal con sobremuestreo

Las métricas de evaluación de este modelo presentan un error $RMSE$: 0.3148 y un coeficiente R^2 : -0.1570.

Diseño de red neuronal con submuestreo

Para el desarrollo de este modelo se ha realizado el procedimiento explicado en 3.5, se presenta en la Figura 3.22 la curva de función de pérdidas para este modelo, donde se ha aplicado el submuestreo con la técnica de Cluster Centroids, con un resultado del estado normal (0) de 1871 a 413 datos y de estado de falla (1) de 413 a 413, quedando con los datos balanceados.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM60

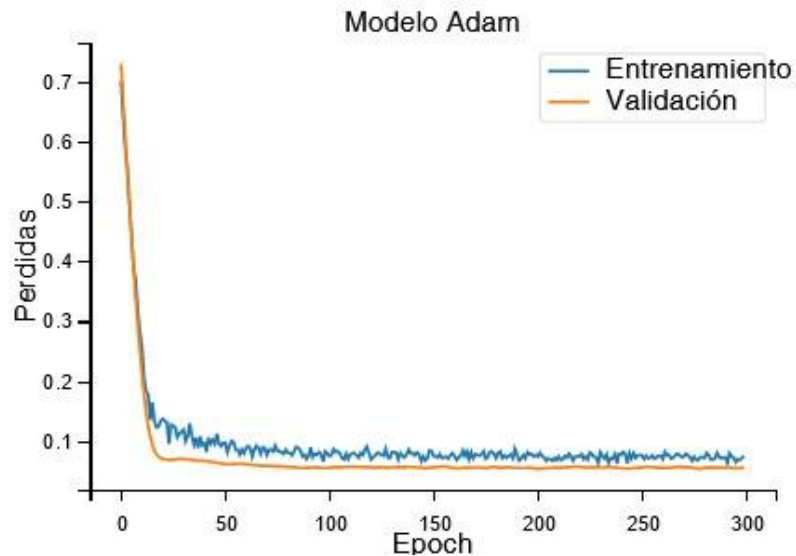


Figura 3.22: Función de pérdidas para envoltorio semanal con Cluster Centroids

Las métricas de evaluación de este modelo presentan un error $RMSE$: 0,4438 y un coeficiente R^2 : -0,7358.

Diseño de red neuronal con sobre y submuestreo

Para el desarrollo de este modelo que combina las técnicas de sobre y sub muestreo, se aplicó el procedimiento explicado en 3.5, donde se ha aplicado en una primera instancia el sobremuestreo con una relación del 60 % entre datos en estado normal y falla, es decir, sin llegar a tener los datos totalmente balanceados, con un resultado parcial del estado normal (0) de 239 a 160 datos y de estado de falla (1) de 81 a 96 datos, posteriormente se ha aplicado el submuestreo, con un resultado del estado normal (0) de 160 a 96 datos y de estado de falla (1) de 96 a 96 datos, quedando con los datos balanceados, la Figura 3.23 presenta la curva de función de pérdidas para este modelo.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM61

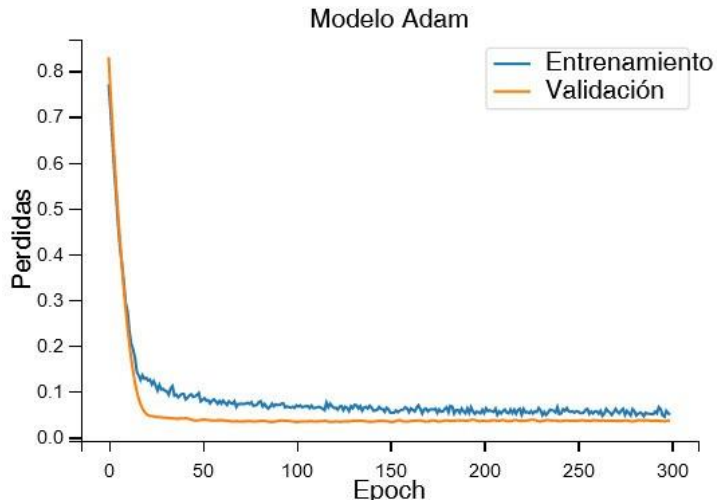


Figura 3.23: Función de pérdidas para envoltorio semanal con técnicas combinadas

Las métricas de evaluación de este modelo presentan un error $RMSE$: 0.2337 y un coeficiente R^2 : -0.1699.

3.5.5. Modelado LSTM con envoltorio mensual

Para el caso de la envoltorio mensual se obtuvieron 38 datos en estado 0 (estado normal) y 37 datos en estado 1 (falla), lo que indica que se tiene datos prácticamente balanceados, por lo que se presenta un modelo con los propios datos generados por la envoltorio mensual, sin la aplicación de estas dos técnicas.

Una vez realizado el procedimiento explicado en 3.5, a excepción del balanceo de datos, se presenta en la Figura 3.24 la curva de función de pérdidas para este modelo.

3.5. DISEÑO DE SISTEMA DE PRONÓSTICO MEDIANTE RNN CON LSTM62

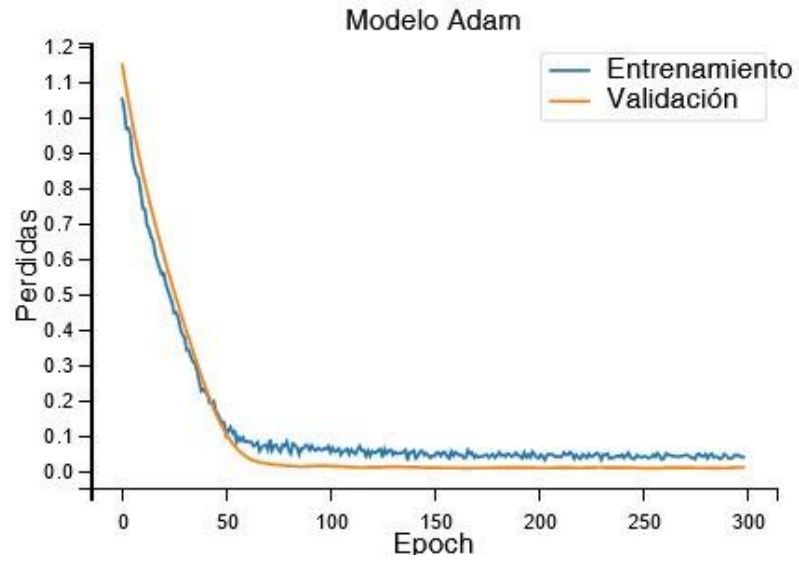


Figura 3.24: Función de pérdidas para envolverte mensual

Las métricas de evaluación de este modelo presentan un error *RMSE*: 0.2222 y un coeficiente R^2 : -0.7780.

Capítulo 4

Análisis de resultados

El capítulo describe el análisis de los resultados obtenidos de las métricas de evaluación y rendimiento de los dos métodos de pronóstico desarrollados: los modelos de regresión, y el sistema mediante redes neuronales recurrentes LSTM, así como se realizan las evaluaciones de los diferentes modelos, y su validación de resultados con las fechas reales de de pronóstico.

4.1. Evaluación de modelos de regresión lineal

Para esta evaluación se tomó los datos de acuerdo a su estado, (0 para los valores de temperatura normal y 1 para los valores de temperatura en falla), y se obtuvieron 3 modelos diferentes, un modelo con datos diarios, uno con datos semanales y el último con datos mensuales, con el objetivo de obtener el pronóstico de la fecha en que ocurra un defecto.

Las métricas de regresión que se han seleccionado para poder evaluar los modelos de regresión lineal son el RMSE, el cual mide la cantidad de error que hay entre dos conjuntos de datos. En otras palabras, compara un valor predicho y un valor observado o conocido. El otro parámetro de evaluación es el R^2 , el cual refleja la calidad del ajuste del modelo a la variable que se desea explicar, este valor da una referencia de cuantos puntos de temperatura están sobre las líneas de las ecuaciones de regresión.

Los valores de RMSE son bajos en los modelos, (Tabla 4.1), pero con un valores de R^2 negativos, lo que indica un modelo de baja calidad, esto debido a que los datos de temperatura del cojinete son muy dispersos y las curvas de regresión de los tres modelos presentados no pueden acoplarse de una forma adecuada para entregar un modelo de alta calidad.

Tabla 4.1: Resumen de métricas de evaluación para regresión

		Regresor Lineal		
		Diario	Semanal	Mensual
Métricas	RMSE	0.0100	0.0300	0.0900
	R^2	-1.5100	-2.7300	-0.9100

4.2. Evaluación de modelos LSTM

Para definir el mejor modelo de pronóstico de falla se realizó una comparación de las métricas de evaluación de los modelos de regresión con los modelos LSTM diarios, semanales y mensuales resultantes del Capítulo 3, las mismas que se presentan en las Tablas: 4.2 (métricas de envolvente diaria), 4.3 (métricas de envolvente semanal) y 4.4 (métricas de envolvente mensual).

Tabla 4.2: Comparación de métricas de evaluación para envoltente diaria

		Regresor			Red Neuronal			Tecnicas para bases desbalanceadas		
		Lineal	Inicial	Optimizada	Smote	Cluster Centroids	Smote / Cluster Centroids			
Métricas	RMSE	0.0100	0.2301	0.2248	0.1254	0.1751	0.2109			
	R ²	-1.5100	0.3797	0.4079	0.5011	0.1249	0.3162			

En la Tabla 4.2 se aprecian las comparaciones de métricas de envoltente diaria con las técnicas de balanceo y el modelo de regresión lineal diario, donde las mejores métricas se producen con la técnica de sobremuestreo SMOTE.

Tabla 4.3: Comparación de métricas de evaluación para envoltente semanal

		Regresor		Tecnicas para bases desbalanceadas		
		Lineal	Smote	Cluster Centroids	Smote / Cluster Centroids	
Métricas	RMSE	0.03	0.3147	0.4437	0.2336	
	R ²	-2.73	-0.157	-0.7357	-0.1699	

En la Tabla 4.3 de comparación de modelos semanales se aprecian que los mejores valores de RMSE y R^2 han sido con la regresión y con la técnica SMOTE, respectivamente, sus valores responden a una cantidad de datos relativamente menor que entrega la envoltente semanal (239 normales y 87 de falla) con respecto a los datos diarios (1871 normales y 413 de falla).

Tabla 4.4: Comparación de métricas de evaluación para envoltente mensual

		Regresor Lineal	Red LSTM
		Métricas	
RMSE	0.0900	0.2222	
R ²	-0.9100	-0.7780	

En la Tabla 4.4 de comparación de modelos mensuales se aprecia que los mejores valores de RMSE y R^2 han sido con la regresión y con la red LSTM, respectivamente, sin aplicación de ninguna técnica de balanceo como se especificó en 3.5.5; sus valores responden a una cantidad de datos menores que los que se obtuvo para la envoltente mensual con respecto a los modelos diarios y semanales.

4.2.1. Evaluación de red neuronal LSTM con envolvente diaria

Las mejores métricas de la red LSTM diaria son con la técnica de sobremuestreo, el RMSE de 0.1254, el cual nos indica una buena precisión con la que el modelo realiza el pronóstico, puesto que los valores bajos del error cuadrático medio dan un mejor ajuste, y el valor de R^2 de 0.5011 que es considerado como valor moderado ([Moore et al., 2016]), siendo el ideal 1, puesto que representa la calidad del modelo.

En la Figura 4.1 se puede apreciar el conjunto de datos de la envolvente diaria con entrenamiento, validación, prueba y pronóstico.

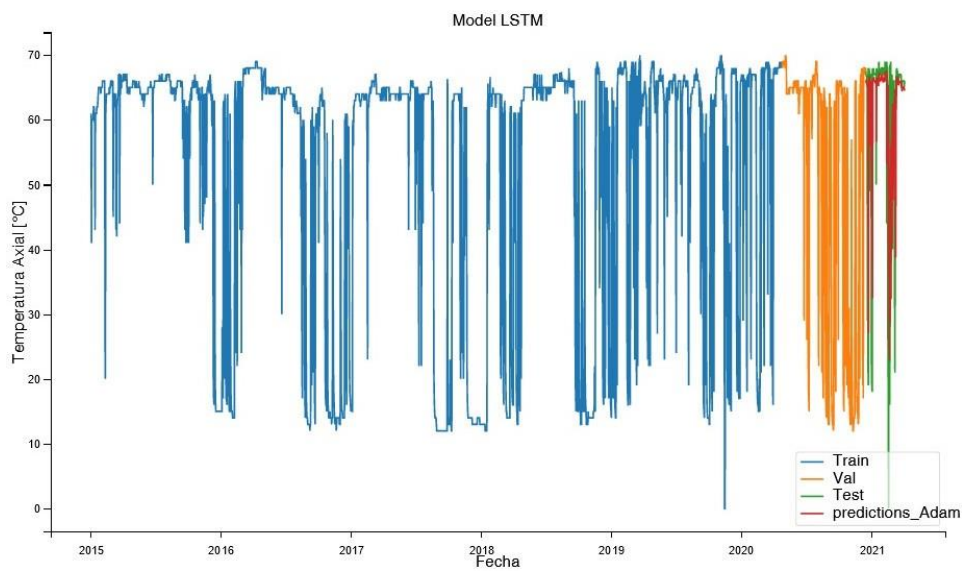


Figura 4.1: Envolvente diaria con entrenamiento, validación, prueba y pronóstico

El sistema realiza un pronóstico de temperatura del cojinete para el 5 de Abril del 2021 de 64.84 °C, Figura 4.2.

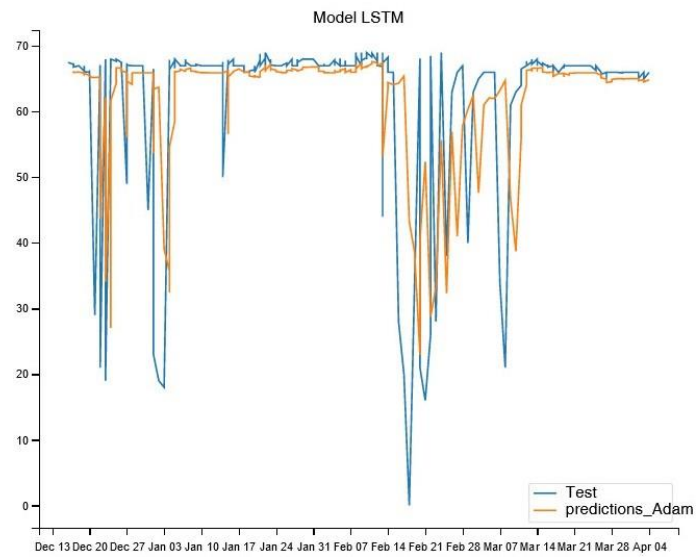


Figura 4.2: Serie temporal de temperatura de cojinete diaria, color azul corresponde a datos reales, color anaranjado corresponde al pronóstico del modelo RNN-LSTM

Para lograr el pronóstico de la red se diseñó un algoritmo free-running, es decir, la salida ha sido realimentada a la entrada, hasta que se ha estimado un valor de temperatura fuera del umbral y de esta manera conocer la fecha posible de fallo, la curva resultante se aprecia en la Figura 4.3.

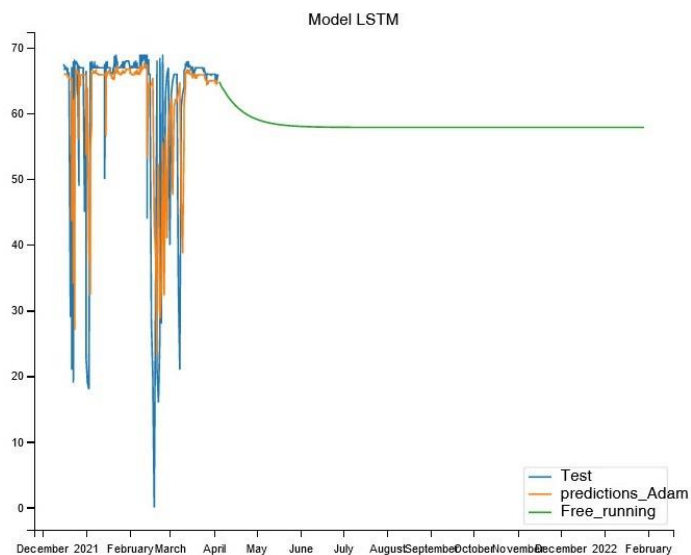


Figura 4.3: Aplicación del algoritmo free-running en envolvente diaria

Para la fecha exacta de pronóstico que entrega el sistema de predicción que es el 05 de abril del 2021, el grupo generador 6 no operó debido a la baja hidrología del mes, siendo el caudal de ingreso a la represa de El Labrado el mínimo obtenido en el histórico desde el año 1998, sin embargo, si se encontró operativa la unidad en días anteriores como el 03 de abril, en el cual cumplió la temperatura de pronóstico de 64°C como se aprecia en la Tabla 4.5. Este valor pronosticado de temperatura forma parte de parámetros normales de operación del generador, por lo que no existió en esta fecha una falla y por lo tanto un mantenimiento.

Tabla 4.5: Valores de temperatura de cojinete de turbina, caudal y potencia de la unidad 6 de Saymirin III-IV presentados el 03-04-2021

HORA	UNIDAD 6				
	TEMPERATURAS			CAUDAL TURB.	POTENCIA (MW)
	C.AXIAL	C.RADIAL	T.C.GEN		
1:00	27,30	26,62	23,45	0,00	0,00
2:00	26,69	25,98	23,10	0,00	0,00
3:00	26,11	25,39	22,80	0,00	0,00
4:00	25,59	24,84	22,60	0,00	0,00
5:00	25,05	24,36	22,16	0,00	0,00
6:00	24,57	23,88	22,01	0,00	0,00
7:00	24,14	23,47	21,60	0,00	0,00
8:00	23,71	23,00	21,42	0,00	0,00
9:00	23,30	22,56	21,10	0,00	0,00
10:00	22,87	22,32	21,06	0,00	0,00
11:00	22,60	22,00	20,60	0,00	0,00
12:00	22,30	21,91	20,60	0,00	0,00
13:00	22,21	21,50	20,60	0,00	0,00
14:00	47,70	44,47	40,08	1,45	2,67
15:00	60,99	52,10	48,98	1,70	3,05
16:00	63,54	53,17	49,60	1,70	3,05
17:00	64,36	53,64	49,60	1,70	3,05
18:00	64,91	53,80	49,60	1,70	3,05
19:00	65,02	53,80	49,60	1,70	3,05
20:00	65,28	53,80	49,60	1,70	3,05
21:00	65,70	53,80	49,49	1,70	3,05
22:00	65,68	53,80	49,23	1,70	3,05
23:00	65,70	53,80	49,10	1,70	3,05
24:00	65,70	53,80	49,10	1,70	3,06

4.2.2. Evaluación de red neuronal LSTM con envoltente semanal

Las métricas de evaluación de este modelo son un RMSE de 0.3147 el cual es un valor de error bajo, pero con un R^2 de -0.157 el cual indica un modelo débil, debido a la baja cantidad de datos que se dispone para el proceso con respecto a los datos diarios.

En la Figura 4.4, se puede apreciar el conjunto de datos de la envoltente semanal con entrenamiento, validación, prueba y pronóstico.

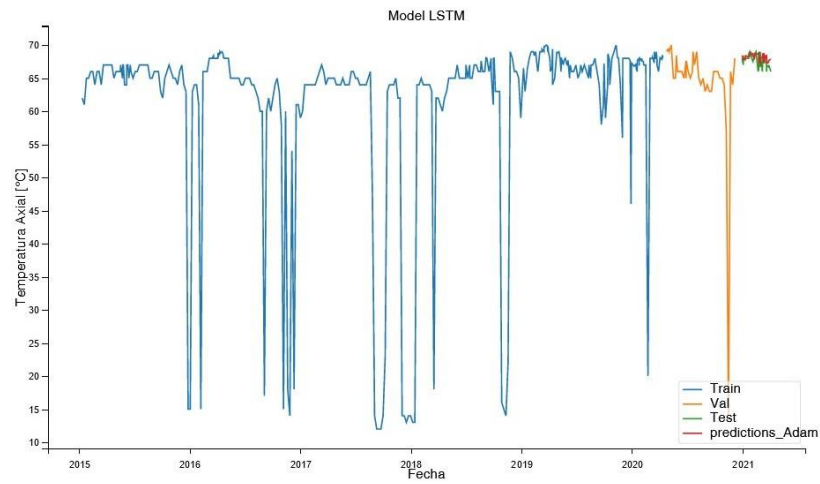


Figura 4.4: Envoltente semanal con entrenamiento, validación, prueba y pronóstico

El sistema realiza un pronóstico de falla del cojinete para el 9 de Mayo del 2021 con una temperatura de 68.54 °C, Figura 4.5.

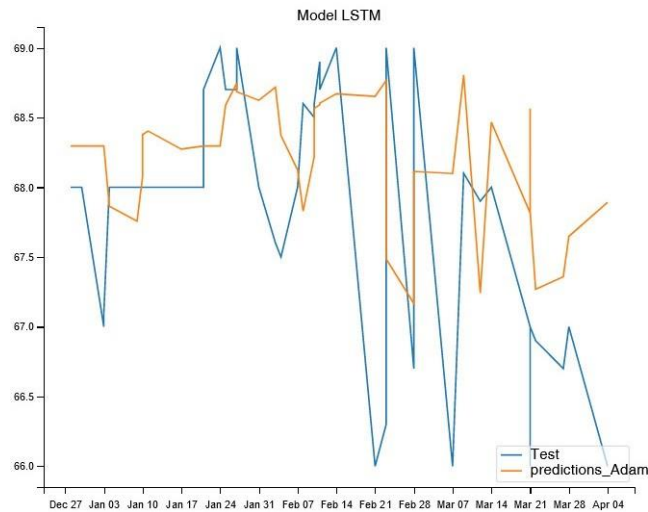


Figura 4.5: Serie temporal de temperatura de cojinete semanal, color azul corresponde a datos reales, color anaranjado corresponde al pronóstico del modelo RNN-LSTM

Para lograr el pronóstico de la red se diseñó un algoritmo free-running, es decir, la salida ha sido realimentada a la entrada, hasta que se ha estimado un valor de temperatura fuera del umbral y de esta manera conocer la fecha posible de fallo, la curva resultante se aprecia en la Figura 4.6.

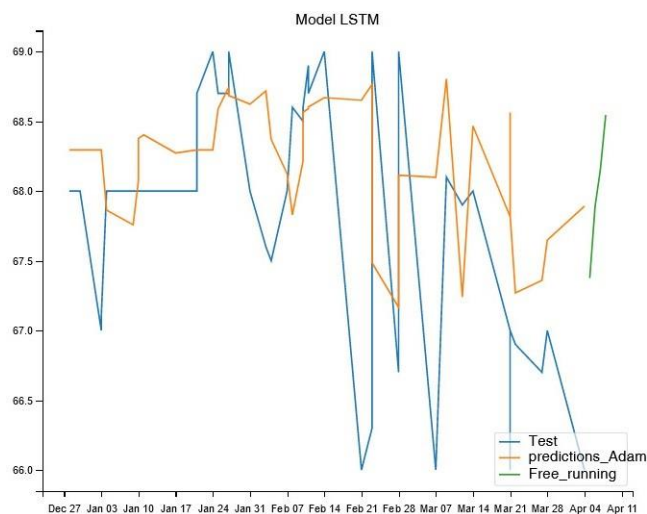


Figura 4.6: Aplicación del algoritmo free-running en envolvente semanal

4.2.3. Evaluación de red neuronal LSTM con envolvente mensual

Las métricas de evaluación de este modelo son un RMSE de 0.2222 el cual es un valor de error bajo, pero con un R^2 de -0.7780. Este modelo resulta ser el de menor calidad con respecto al semanal y diario, debido a la poca cantidad de datos (38 normales y 37 de falla) que se obtuvo como envolvente mensual, sin aplicación de balanceo.

En la Figura 4.7, se puede apreciar el conjunto de datos de la envolvente mensual con entrenamiento, validación, prueba y pronóstico.

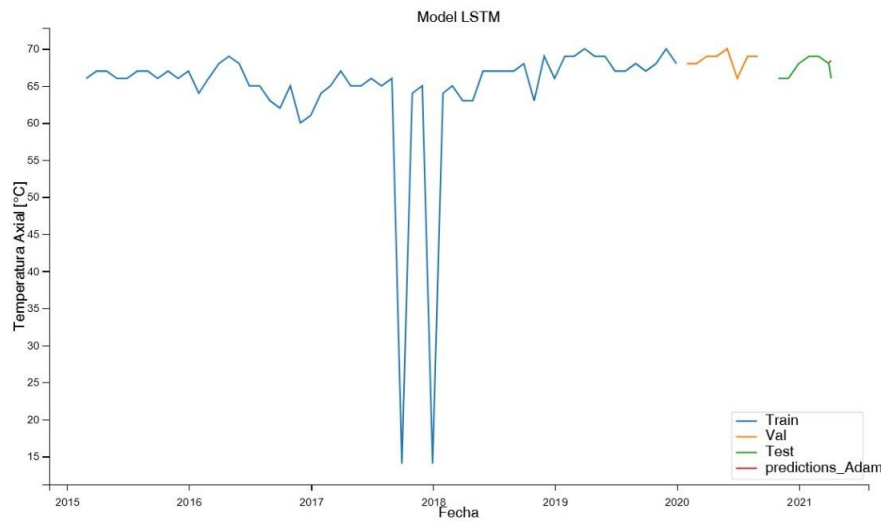


Figura 4.7: Envolvente mensual con entrenamiento, validación, prueba y pronóstico

El sistema realiza un pronóstico de falla del cojinete para el mes de Mayo del 2021 con una temperatura de 68.86 °C, Figura 4.8.

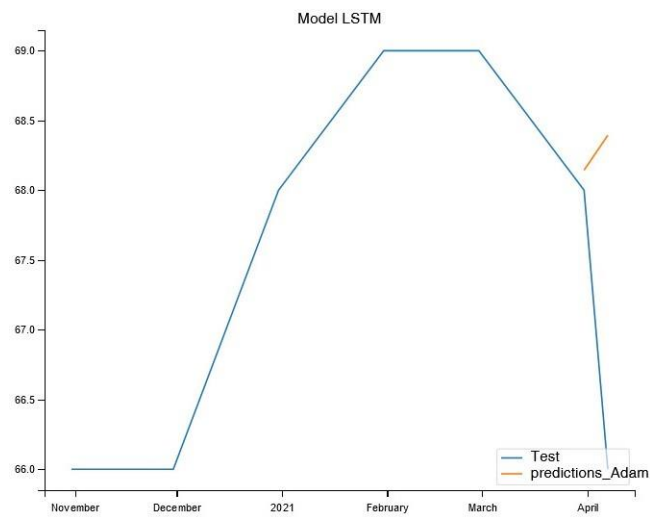


Figura 4.8: Serie temporal de temperatura de cojinete mensual, color azul corresponde a datos reales, color anaranjado corresponde al pronóstico del modelo RNN-LSTM

Para lograr el pronóstico de la red se diseñó un algoritmo free-running, es decir, la salida ha sido realimentada a la entrada, hasta que se ha estimado un valor de temperatura fuera del umbral y de esta manera conocer la fecha posible de fallo, la curva resultante se aprecia en la Figura 4.9.

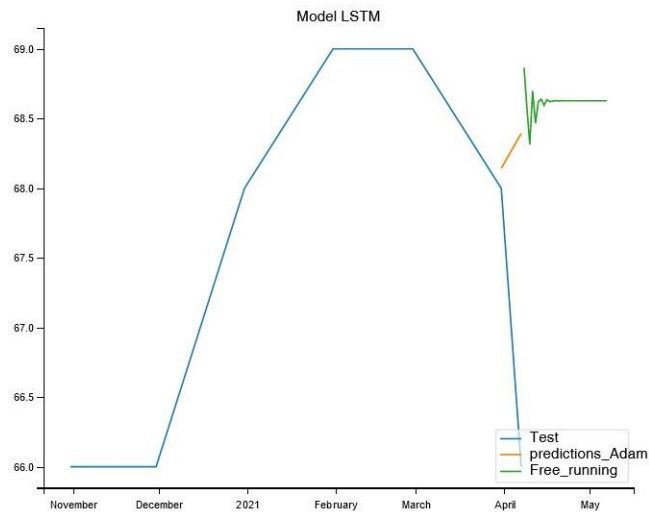


Figura 4.9: Aplicación del algoritmo free-running en envolvente mensual

El mejor modelo de pronóstico es el logrado por medio de la red LSTM con la técnica de sobremuestreo SMOTE de envolvente diaria con una métrica RMSE de 0.1254 y un R^2 de 0.5011, siendo esta muy superior a lo obtenido por medio de los modelos de regresión lineal. La capacidad de generalización de las redes neuronales es muy superior a los resultados obtenidos por la regresión.

La Tabla 4.6 presenta las métricas de evaluación de todos los modelos que se han desarrollado en este proyecto, siendo notable que los mejores resultados se presentan por medio de las redes neuronales.

Tabla 4.6: Métricas de evaluación para los modelos de pronóstico desarrollados

		Tipo / Técnica	Métricas	
			RMSE	R2
Regresión Lineal		Diaria	0,010	-1,51
		Semanal	0,030	-2,73
		Mensual	0,090	-0,91
Red Neuronal LSTM	Envolvente diaria	SMOTE	0,125	0,501
		Cluster Centroids	0,175	0,124
		SMOTE/ Cluster Centroid	0,210	0,316
	Envolvente semanal	SMOTE	0,314	-0,157
		Cluster Centroids	0,443	-0,735
		SMOTE / Cluster Centroid	0,233	-0,169
	Envolvente mensual	Modelo sin balancear	0,222	-0,778

Capítulo 5

Conclusiones y trabajo futuro

Luego de haber realizado todas las pruebas pertinentes a los modelos de predicción y las técnicas de balanceo de clases desarrolladas en este trabajo, se puede elaborar diversas deducciones que se detallan a continuación:

- Para realizar un pronóstico de la temperatura de cojinete es necesario disponer de información confiable y suficiente para desarrollar un análisis detallado de la toda la información disponible que se relacione con el estado de la máquina, fechas y horas. Debido al alto grado de complejidad del análisis, la eficiencia de la predicción depende directamente de la robustez del modelo y las técnicas a utilizar para el mejoramiento de su precisión.
- La calidad y veracidad de los datos, es información fundamental para el desarrollo de modelos de predicción adecuados.
- El modelo de regresión lineal es fácil de entender y explicar, además se puede modelar rápidamente, sin embargo, no permite capturar relaciones no lineales de entrada, por lo que es difícil obtener una recta que represente todas las mediciones.
- Al implementar el algoritmo LSTM con optimizador Adam, se ha observado que tiene una mejor representación de los datos que los métodos de regresión lineal, pero las métricas son poco favorables; como se observa en el Capítulo 4, Tabla 4.2, lo que hace que se descarte en una primera instancia como método para la predicción. En este sentido se tiene un coeficiente de determinación R^2 de 0,3797 que es un valor insatisfactorio.

- En el Capítulo 3, se llevó a cabo el procedimiento para la selección del número de épocas, iteraciones y neuronas a ser utilizadas, se observó que el número de neuronas en la capa oculta varía dependiendo al número de entradas por lo que se debe realizar diferentes entrenamientos para obtener mejor ajuste de la red neuronal. No existe una regla definida para determinar el número de neuronas en cada capa, por lo que el criterio que se utilizó fue el de entrenar varias veces el modelo y analizar los resultados de las métricas obtenidas para cada uno de los modelos de pronóstico resultantes.
- Basándose en dicho criterio en la sección 3.5.1 del Capítulo 3, se determinó que el número de neuronas para la capa LSTM sea de cuatro y para la capa densa de dieciséis; como se observa en los resultados presentados en la Tabla 3.4; si se selecciona un número incorrecto de neuronas en la capa oculta se sobrestima o se subestima los valores de la red proporcionando pronósticos erróneos, estos resultados se visualizan de mejor manera en el mapa de calor de la Figura 3.15. En este contexto, en la Figura 3.16 se puede evidenciar la mejora obtenida en relación al modelo inicial de la red LSTM (Figura 3.13).
- Debido a que no existe una metodología establecida que mencione como seleccionar el porcentaje de datos de entrenamiento y validación, se efectuó una serie de ensayos durante el proceso de entrenamiento de la red neuronal para llegar a concluir que: el seleccionar adecuadamente el porcentaje de los datos para el entrenamiento de la red neuronal, es de vital importancia para obtener resultados óptimos en los pronósticos, ya que una mala selección de dichos porcentajes altera en la operatividad de la red neuronal, el 80 % de los datos de entrada de la red neuronal se utilizó para realizar el entrenamiento de la red, se seleccionó dicho porcentaje debido a que la red se ajusta iteración a iteración en base al error de entrenamiento, 10 % de los datos de entrada a la red neuronal para realizar la validación y el restante 10 % únicamente para evaluación, es decir este último conjunto no se utilizó para modificar ningún parámetro de la red.
- Es recomendable realizar el escalamiento (normalización) de los datos que serán utilizados en la predicción, es decir transformarlos en valores dentro del rango de cero a uno, para facilitar el procesamiento y análisis de los mismos y por consecuencia obtener mejores resultados disminuyendo el porcentaje en las medidas de error en la predicción.

- De acuerdo a lo realizado en el Capítulo 3 y en función de los resultados mostrados en las Tablas 3.1 a la 3.7, se ha analizado el comportamiento que adopta el pronosticador si se modifican los parámetros de su construcción, en este caso se pudo observar que el número óptimo de épocas e iteraciones son doscientas y veinte respectivamente, de igual manera la cantidad óptima de neuronas y dropout en la capa LSTM fueron cuatro y cuarenta por ciento, mientras que en la capa densa se tuvo dieciséis neuronas con un dropout de veinte por ciento como se resume en la Tabla 3.8.
- Al modificar el optimizador de la red LSTM y de acuerdo a las métricas presentadas en las Tablas 3.1 a la 3.3, se pudo determinar que con el optimizador Adam se obtuvieron mejores métricas en comparación con los optimizadores RMSPROP y Adagrad, llegando a tener un error de 0,2280, siendo el mejor de los 3 optimizadores; sin embargo, por la naturaleza de los datos se presentó el problema de desbalanceo de clases, lo cual hizo necesario aplicar técnicas para tratar este tipo de información.
- Las técnicas de sub muestreo propuestas en este trabajo para balancear las clases no mejoraron los resultados obtenidos en el algoritmo LSTM, al contrario, las métricas del algoritmo se vieron afectadas, no así la técnica SMOTE de sobremuestreo, que presentó una disminución en el error y una mejora notoria en el coeficiente de determinación R^2 . (Tabla 4.2).
- Al aplicar el algoritmo free-running en el mejor modelo con la técnica de sobremuestreo se puede observar que la curva de pronóstico, Figura 4.3 trabaja mejor en rangos cortos ya que se pudo observar que a partir de la iteración 131 el valor de la temperatura se mantiene constante; sin embargo, este algoritmo permitió identificar que el pronóstico realizado de la serie temporal no converge a un modelo trivial como se puede observar en la Figura 4.9.
- Con el desarrollo del presente trabajo se pudo observar que el rendimiento de los diferentes algoritmos depende en gran medida del tamaño y la estructura de los datos.
- Para aplicaciones futuras de pronóstico, se recomienda tomar en cuenta otras variables adicionales que puedan ser sensibles en el pronóstico, como es el caso de vibración que actualmente no se cuenta con estas

mediciones; esto con el fin de que el modelo sea más afinado y permita seguir la tendencia real de la unidad de generación.

- Como trabajo futuro queda realizar un análisis en otros equipos y sistemas de las centrales, como es el caso de análisis de fallas en transformadores de potencia de centrales de generación, para la interpretación de los diferentes niveles de gases que contiene el aceite dieléctrico, lo que aportaría información para la toma de decisiones en los programas de mantenimiento que se realizan a estos equipos.

Bibliografía

- AENOR. Norma Española Terminología del Mantenimiento, 2011.
- C. C. Aggarwal. *Neural Networks and Deep Learning*. NY, USA, 2018. ISBN 9783319944623. doi: <https://doi.org/10.1007/978-3-319-94463-0>.
- V. L. G. Ajuech. *Mantenimiento: técnicas y aplicaciones industriales*. 2017.
- L. Aldaz, B. Eguía, and L. Urcola. Diseño conceptual y Normalización de datos. pages 1–21, 2009.
- R. Arno, N. Dowling, R. S. Pe, H. P. Cfs, and E. Y. P. Mcf. Equipment Failure Characteristics and RCM for Optimizing Maintenance Cost. *IEEE Transactions on Industry Applications* (Volume: 52, Issue: 2, March-April 2016), 9994(c):1–8, 2015. doi: 10.1109/TIA.2015.2483700. URL <https://ieeexplore.ieee.org/document/7283613>.
- R. Arunthavanathan, F. Khan, S. Ahmed, and S. Imtiaz. A deep learning model for process fault prognosis. *Process Safety and Environmental Protection*, 154:467–479, 2021. ISSN 0957-5820. doi: 10.1016/j.psep.2021.08.022. URL <https://doi.org/10.1016/j.psep.2021.08.022>.
- L. A. Baumgart, E. J. Bass, S. Member, J. D. Voss, and J. A. Lyman. Prediction of Bearing Degradation Trend based on LSTM. *IEEE Transactions on Human-Machine Systems*, 45(6):773–781, 2015. doi: 10.1109/THMS.2015.2459382.
- S. Bianchi, R. Paggi, G. L. Mariotti, and F. Leccese. Why and When Must The Preventive Maintenance be Performed ? pages 221–226, 2014.
- T. Boukra and A. Lebaroud. Identifying New Prognostic Features for Remaining Useful Life Prediction. *2014 16th International Power Electronics and Motion Control Conference and Exposition*, pages

- 1216–1221, 2014. doi: 10.1109/EPEPEMC.2014.6980677. URL <https://ieeexplore.ieee.org/document/6980677>.
- A. Cachada, J. Barbosa, P. Leitão, C. A. S. Geraldés, L. Deusdado, J. Costa, C. Teixeira, J. Teixeira, A. H. J. Moreira, P. M. Moreira, and L. Romero. Maintenance 4 . 0 : Intelligent and Predictive Maintenance System Architecture. *2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 139–146, 2018. doi: 10.1109/ETFA.2018.8502489. URL <https://ieeexplore.ieee.org/document/8502489>.
- W. Caesarendra, A. Widodo, P. H. Thom, B.-s. Yang, and J. D. Setiawan. Combined Probability Approach and Indirect Data-Driven Method for Bearing Degradation Prognostics. *IEEE Transactions on Reliability (Volume: 60, Issue: 1, March 2011)*, 60(1):14–20, 2011. doi: 10.1109/TR.2011.2104716.
- J. Calderón, J. Cadavid, and D. Ovalle. NEURAL NETWORK FOR FAULT CLASSIFICATION IN TRANSMISSION LINES BASED ON OSCILOPERTUBOGRAPHY RECORDS. pages 99–107, 2008.
- G. C. Canavos. *Applied Probability and Statistical Methods*. McGraw Hill. U.S.A., México, 1998. ISBN 968-451-856-0.
- L. Capelo. *Beginning Application Development with TensorFlow and Keras*. 2018.
- A. A. Castañares. Aplicación del dropout a la cuantificación de la incertidumbre en redes neuronales. 2020.
- S.-t. Chang. Lifetime Prediction for Bearings in Induction Motor. *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, pages 467–471, 2019. doi: 10.1109/ICPHYS.2019.8780366.
- S. Chatterjee and A. S. Hadi. *Regression Analysis by Example*. A John Wiley and Sons, Ltd., Publication Francq., 5 edition, 2012. ISBN 978-0-470-90584-5.
- Z. Chen and Y. Liu. Mechanical State Prediction Based on LSTM Neural Netwok. *2017 36th Chinese Control Conference (CCC)*, pages 3876–3881, 2017. doi: 10.23919/ChiCC.2017.8027963. URL <https://ieeexplore.ieee.org/document/8027963>.

- J. Cortés, N. Bielsa, P. Muñoz, Á. Ruiz, and J. A. González. *Bioestadística para no Estadísticos*. Universidad Politecnica de Catalunya, 2015.
- I. Delgado. *Estudio de los Sistemas de Regulación de Velocidad y Turbinas Hidroeléctricas y su influencia en la Estabilidad Transitoria de Sistemas Eléctricos de Potencia*. PhD thesis, Escuela Politécnica Nacional, 1997.
- J. Deutsch and D. He. Using Deep Learning-Based Approach to Predict Remaining Useful Life of Rotating Components. *IEEE Transactions on Systems, Man, and Cybernetics: Systems (Volume: 48, Issue: 1, Jan. 2018)*, pages 1–10, 2017. doi: 10.1109/TSMC.2017.2697842.
- H. Ding, L. Yang, Z. Cheng, and Z. Yang. A remaining useful life prediction method for bearing based on deep neural networks. *Measurement*, 172(October 2020):108878, 2021. ISSN 0263-2241. doi: 10.1016/j.measurement.2020.108878. URL <https://doi.org/10.1016/j.measurement.2020.108878>.
- J. Duchi, E. Hazan, and Y. Singer. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *Journal of Machine Learning Research 12 (2011) 2121-2159*, 12:2121–2159, 2011.
- É. Espejo and H. Hernández. *Análisis de fallas de estructuras y elementos mecánicos*. Editorial Universidad Nacional de Colombia., 2017. ISBN 9789587838619, 9789587830224. URL <https://bibliotecas.ups.edu.ec:3488/es/ereader/bibliotecaups/130007?page=747>.
- C. Francq, J. Zakoian, and J. Zakoian. *GARCH Models Structure , Statistical Inference and Financial Applications*. Number 2010. A John Wiley and Sons, Ltd., Publication Francq., 2021.
- Q. Fu, H. Wang, J. Zhao, and X. Yan. A Maintenance-prediction Method for Aircraft Engines using Generative Adversarial Networks. *2019 IEEE 5th International Conference on Computer and Communications (ICCC)*, pages 225–229, 2019. doi: 10.1109/ICCC47050.2019.9064184. URL <https://ieeexplore.ieee.org/document/9064184>.
- D. Gallará, I. y Pontelli. *Mantenimiento Industrial*. 2020.
- W. J. Gil Gonzalez, J. J. Mora Flórez, and S. M. Pérez Londoño. Análisis del procesamiento de los datos de entrada para un localizador de fallas en sistemas de distribución. 18(41):64–75, 2014.

- L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing*, 2017. ISSN 0925-2312. doi: 10.1016/j.neucom.2017.02.045. URL <http://dx.doi.org/10.1016/j.neucom.2017.02.045>.
- J. F. Sanz Osorio. *Energía hidroeléctrica (2a. ed.)*. Zaragoza, 2016.
- A. Kattan, R. Abdullah, and Z. W. E. Geem. Artificial Neural Network Training and Software Implementation Techniques. pages 1–2, 2011.
- X.-h. Le, H. V. Ho, G. Lee, and S. Jung. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. 2019.
- B.E. López Porrero. *Limpieza de datos*. Editorial Feijóo, Santa Clara, Cuba, 2009. URL <https://bibliotecas.ups.edu.ec:3488/es/ereader/bibliotecaups/71744?page=14>.
- T. H. Loutas, D. Roulias, and G. Georgoulas. Remaining Useful Life Estimation in Rolling Bearings Utilizing Data-Driven Probabilistic E-Support Vectors Regression. *IEEE Transactions on Reliability (Volume: 62, Issue: 4, Dec. 2013)*, 62(4):821–832, 2013. doi: 10.1109/TR.2013.2285318. URL <https://ieeexplore.ieee.org/document/6645455>.
- Z. Lt, S. Zhang, and W. Shi. An Integrated Method to Detect the Incipient Degradation of Bearings by Vibration Analysis and Feature Extraction. *2016 International Conference on Condition Monitoring and Diagnosis (CMD)*, 2016. doi: 10.1109/CMD.2016.7757914. URL <https://ieeexplore.ieee.org/document/7757914>.
- J. Ma. State-based Time-Series Analysis and Prediction. pages 227–232, 2007. doi: 10.1109/SNPD.2007.363.
- S. Marble and B. P. Morton. Predicting the Remaining Life of Propulsion System Bearings. *2006 IEEE Aerospace Conference*, 2006. doi: 10.1109/AERO.2006.1656121. URL <https://ieeexplore.ieee.org/document/1656121>.
- P. Méndez and J.T. *Inteligencia artificial: métodos, técnicas y aplicaciones*. España, 2008. URL <https://bibliotecas.ups.edu.ec:3488/es/ereader/bibliotecaups/50116?page=699>.
- Mike Bernico. *Deep Learning Quick Reference*. 2018. ISBN 9781788837996.

- D. C. Montgomery, E. A. Peck, and G. G. Vining. *Introduction to Linear Regression Analysis*. A John Wiley and Sons, Ltd., Publication Francq,, 5 edition, 2012. ISBN 978-0-470-54281-1.
- D. S. Moore, W. Notz, and M. A. Fligner. *The Basic Practice of Statistics*, volume 3. W.H. Freeman Co., Beaverton, 7th editio edition, 2016. ISBN 9781464142536. URL <https://www.proquest.com/other-sources/basic-practice-statistics-7th-editiononline/%0Adocview/1787995450/se-2?accountid=32861>.
- G. Niu, S. Tang, and B. Zhang. Machine Condition Prediction based on Long Short Term Memory and Particle Filtering. *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, 1:5942–5947, 2018.
- P. Poor, J. Basl, and D. Zenisek. Predictive Maintenance 4 . 0 as next evolution step in industrial maintenance development. *2019 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, pages 245–253, 2019. doi: 10.23919/SCSE.2019.8842659. URL <https://ieeexplore.ieee.org/document/8842659>.
- R. Ortiz Flórez. *Pequeñas Centrales Hidroeléctricas*. 2011. ISBN 9781449278335, 9789588675992.
- N. N. Ravi, S. M. Drus, and P. S. Krishnan. Data Mining Techniques for Transformer Failure Prediction Model : A Systematic Literature Review. *2019 IEEE 9th Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, pages 305–309, 2019. doi: 10.1109/ISCAIE.2019.8743987. URL <https://ieeexplore.ieee.org/document/8743987>.
- W. Shuqing. Machine Health Monitoring and Prognostication Via Vibration Information. *Sixth International Conference on Intelligent Systems Design and Applications*, pages 0–4, 2006. doi: 10.1109/ISDA.2006.188. URL <https://ieeexplore.ieee.org/document/4021555>.
- D. Siegel, J. Lee, and C. Ly. Methodology and Framework for Predicting Rolling Element Helicopter Bearing Failure. *2011 IEEE Conference on Prognostics and Health Management*, 2011. doi: 10.1109/ICPHM.2011.6024339.
- A. Soria Lorente, E. R. Moreno Roque, and J. B. Martí Zamora. *Polinomios ortogonales clásicos de una variable discreta , su historia , extensiones y aplicaciones*, volume 35. 2014.

- P. Taylor, N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan. Residual-life distributions from component degradation signals : A Bayesian approach Residual-life distributions from component degradation signals : A Bayesian approach. (July 2013):37–41, 2007. doi: 10.1080/07408170590929018. URL <http://dx.doi.org/10.1080/07408170590929018>.
- E. Uriel. *Contraste de hipótesis en el modelo de regresión múltiple*. Universidad de Valencia., 1998.
- S. V. Verdú. *Aplicación de un modelo de red neuronal no supervisado a la clasificación de consumidores eléctricos*. San Vicente (Alicante): ECU, 2013, 2013. ISBN 9788499486758.
- H. Vildan, M. Taskiran, and N. Kahraman. LSTM and WaveNet Implementation for Predictive Maintenance of Turbofan Engines. *IEEE 20th International Symposium on Computational Intelligence and Informatics*, pages 151–156, 2020.
- D. Wang and K.-l. Tsui. Statistical Modeling of Bearing Degradation Signals. *IEEE Transactions on Reliability (Volume: 66, Issue: 4, Dec. 2017)*, pages 1–14, 2017. doi: 10.1109/TR.2017.2739126. URL <https://ieeexplore.ieee.org/document/8023771>.
- Q. Wang, B. Wei, and J. Liu. Data-driven incipient fault prediction for non- stationary and non-linear rotating systems : methodology , model construction and application. XX, 2020. doi: 10.1109/ACCESS.2020.3032445.
- R.Wang, H. Jiang, X. Li, and S. Liu. A Reinforcement Neural Architecture Search Method for Rolling Bearing Fault Diagnosis. *Measurement*, page 107417, 2019. ISSN 0263-2241. doi: 10.1016/j.measurement.2019.107417. URL <https://doi.org/10.1016/j.measurement.2019.107417>.
- S. Weisberg. *Applied Linear Regression*. John Wiley and Sons, Inc., Hoboken, New Jersey, fourth edi edition, 2014. ISBN 9781118386088.
- Y. Xiao. Hybrid LSTM Neural Network for Short-Term Traffic Flow Prediction. 2019. doi: 10.3390/info10030105.
- Z. Xu, C. Li, and Y. Yang. Fault Diagnosis of Rolling Bearings Using an Improved Multi-Scale Convolutional Neural Network with Feature Attention mechanism. *ISA Transactions*, 2020. ISSN 0019-0578. doi: 10.

- 1016/j.isatra.2020.10.054. URL <https://doi.org/10.1016/j.isatra.2020.10.054>.
- Q. Ye, X. Yang, C. Chen, and J. Wang. River Water Quality Parameters Prediction Method Based on LSTM-RNN Model. *2019 Chinese Control And Decision Conference (CCDC)*, pages 3024–3028, 2019.
- G. Yuhai, L. Shuo, and H. Linfeng. Research on Failure Prediction Using DBN and LSTM Neural Network. *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pages 1705–1709, 2018.
- R. Zaheer and H. Shaziya. A Study of the Optimization Algorithms in Deep Learning. *2019 Third International Conference on Inventive Systems and Control (ICISC)*, (August):536–539, 2020. doi: 10.1109/ICISC44355.2019.9036442. URL <https://ieeexplore.ieee.org/document/9036442>.
- T. Zhang and X. You. Improvement of the training and normalization method of artificial neural network in the prediction of indoor environment. *9th International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC) and the 3rd International Conference on Building Energy and Environment (COBEE)*, 121:1245–1251, 2015. ISSN 1877-7058. doi: 10.1016/j.proeng.2015.09.152. URL <http://dx.doi.org/10.1016/j.proeng.2015.09.152>.
- Y. Zhang, T. Zhou, X. Huang, L. Cao, and Q. Zhou. Fault diagnosis of rotating machinery based on recurrent neural networks. *Measurement*, 171(October 2020):108774, 2021. ISSN 0263-2241. doi: 10.1016/j.measurement.2020.108774. URL <https://doi.org/10.1016/j.measurement.2020.108774>.
- Z. Zhu, H. Wang, Z. Liu, and S. Meng. Fault Diagnosis of Wheelset Bearings using Deep Bidirectional Long Short-term Memory Network. *2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, pages 1–7, 2019. doi: 10.1109/PHM-Qingdao46334.2019.8942870. URL <https://ieeexplore.ieee.org/document/8942870>.

Glosario

- **PdM:** Enfoques de predicción de eventos
- **RUL:** Vida útil restante
- **RNN:** Redes Neuronales Recurrentes
- **LSTM:** Long Short Term Memory
- **FFT:** Transformada rápida de fourier
- **STFT:** Transformada de fourier de corta duración
- **SVR:** Algoritmo de vectores de regresión
- **LR:** Regresión logística
- **SCADA:** Supervisión, control y adquisición de datos
- **BD:** Base de datos
- **MC:** Mínimos cuadrados
- **BM:** Algoritmo de retropropagación
- **MLP:** Perceptrón multicapa
- **GD:** Gradiente descendiente
- **Adagrad:** Adaptative gradient algorithm
- **RMSPROP:** Root mean square propagation
- **Adam:** Adaptative moment optimization
- **SMOTE:** Synthetic minority oversampling technique