

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Ciencias de la Tierra

Modelamiento predictivo secuencial para concentraciones de metano en minería subterránea de carbón vía redes neuronales recurrentes.

PROYECTO INTEGRADOR

Previo la obtención del Título de:

Ingeniero en Minas

Presentado por:

Carlos Alberto Añazco Toro

GUAYAQUIL - ECUADOR

Año: 2022

DEDICATORIA

“La fortaleza de un hombre yace en
sus raíces”

Agradezco a Dios por tener una gran mujer como madre. Este proyecto es posible gracias a todo el tiempo dedicado por ella a la crianza y educación de mis amados hermanos y yo, no seríamos nadie sin ella. De igual forma, agradezco a mi padre presente en nuestra educación.

Gracias a mi esposa e hijo que forman hoy en día un pilar fundamental que sostiene mi diario caminar y con amor dedico todo el esfuerzo realizado en estos largos años a mi abuelita Rosa y mis ángeles Juan José y Silvia.

AGRADECIMIENTOS

Mis más sinceros agradecimientos a los tutores y maestros que formaron mi vida académica de una excelente manera, ellos han hecho este proyecto posible. El mismo cuenta con un gran alcance para la automatización de procesos y servicios mineros, aportando estos conocimientos a la industria minería ecuatoriana con gran orgullo y disposición.

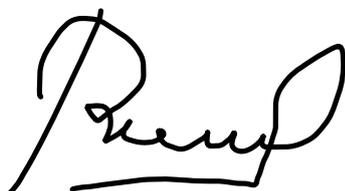
Doctores:

PhD. Samantha Jiménez

PhD (c). Jorge Loy

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, me corresponde conforme al reglamento de propiedad intelectual de la institución; *Carlos Alberto Añezco Toro*, doy mi consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”

A handwritten signature in black ink, appearing to read 'Añezco Toro', with a large initial 'A' and a stylized 'T'.

Carlos Alberto Añezco
Toro

EVALUADORES

A handwritten signature in blue ink, reading "Samantha Jiménez Oyola", is centered on the page. The signature is written in a cursive style with a large initial 'S'. Below the signature is a horizontal dotted line.

Ph.D. Samantha Jiménez Oyola

TUTOR Y PROFESOR DE LA MATERIA

RESUMEN

El modelamiento predictivo y su uso en distintos campos de la investigación ha optimizado los procesos ingenieriles en la minería y otras industrias. El objetivo de este proyecto es implementar un modelo predictivo capaz de evaluar y predecir la concentración de metano en labores mineras subterráneas de carbón, además de identificar las condiciones internas de la mina que puedan favorecer la acumulación de gas metano en los frentes de trabajo. Para este propósito, se utilizó el lenguaje de programación Python, con la finalidad de optimizar el preprocesamiento de los datos y eliminar anomalías en la data. La evaluación y mejora continua del modelo se realizó por funciones internas de programación. La arquitectura de la red neuronal procesó la información de 34 variables de entrada y predijo series temporales de 3 sensores objetivos que captaron altas concentraciones de metano. Se encontró que los modelos neuronales bidireccionales aplicados en este estudio arrojaron excelentes resultados en comparación con las redes recurrentes normales. Se comparó los valores reales con los valores obtenidos por el modelo, observando una correlación de exactitud del 99.50%. En este sentido, el modelo de predicción que mejor se adapta a la data obtenida en el interior de mina es el BILSTM.

Palabras Clave: Redes neuronales recurrentes, Minería subterránea de carbón, Sistema de alerta temprana, Metano, Modelos bidireccionales, Seguridad minera.

ABSTRACT

Predictive modeling is used in different fields of research optimized engineering processes in mining and other industries. The objective of this project is to implement a predictive model capable of evaluating and predicting concentration of methane in underground coal mining operations. It's important to identify the mine internal conditions could be favor the methane accumulation gas in work fronts. For this purpose, the Python programming language was used to optimize the preprocessing data and eliminate anomalies present in data set. The evaluation and continuous improvement of the model was carried out by internal programming functions. The neuronal network architecture processed information from 34 input variables and predicted time series from 3 target sensors that captured high methane concentrations. The bidirectional neuronal model applied in this study were found to yield excellent results compared to normal recurrent networks. The real values were compared with the values obtained by the model, observing an accuracy correlation of 99.50%. In this sense, the prediction model that best adapts to the data collected inside the mine is the BILSTM.

Keywords: Recurrent Neural Networks, Underground Coal Mining, Warning System, Methane, Bidirectional Models, Security Mining.

ÍNDICE GENERAL

EVALUADORES.....	5
RESUMEN.....	I
<i>ABSTRACT</i>	II
ÍNDICE GENERAL.....	III
ABREVIATURAS	V
SIMBOLOGÍA	VI
ÍNDICE DE FIGURAS.....	VII
ÍNDICE DE TABLAS	VIII
CAPÍTULO 1.....	9
1. Introducción	9
1.1 Descripción del problema	11
1.2 Justificación Del Problema.....	12
1.3 Objetivos.....	12
1.3.1 Objetivo General	12
1.3.2 Objetivos Específicos	12
1.4 Marco teórico	12
1.4.1 Características del sitio de estudio	12
1.4.1.1 Ubicación de la Mina	12
1.4.1.2 Geología	13
1.4.2 Descripción de las Operaciones Mineras	14
1.4.3 Aprendizaje Profundo, Herramientas y Aplicación.....	15
1.4.5 Función De Auto correlación	27
CAPÍTULO 2.....	29
2. Metodología	29

2.1 Descripción de los datos.....	30
2.2 Sensores y Características.....	33
2.3 Herramientas y Programa Usado.....	33
2.4 Detección de Anomalías Vía análisis de Componentes Principales.....	34
2.5 Selección de atributos vía eliminación recursiva de funciones (RFE).....	35
2.6 Procesamiento de Ventana Móvil y su Determinación Vía función de Auto correlación.....	35
2.7 Arquitectura de la Red Neuronal.....	36
2.8 Desarrollo de Sistema de Alerta Temprana Vía Redes Neuronales Recurrentes	
37	
CAPÍTULO 3.....	38
3. RESULTADOS.....	38
3.1 Resultados PCA.....	38
3.2 Resultados RFE.....	39
3.3 Función de auto correlación para múltiples sensores de metano.....	40
3.4 Entrenamiento de modelos y optimización.....	42
3.5 Evaluación de modelos predictivos.....	46
CAPÍTULO 4.....	49
4. Conclusiones Y RECOMENDACIONES.....	49
4.1 Conclusiones y trabajo futuro.....	49
4.2 Recomendaciones.....	50
BIBLIOGRAFÍA.....	52
ANEXOS.....	55
Anexo A.....	56
Anexo B.....	57
Anexo C.....	58

ABREVIATURAS

ESPOL	Escuela Superior Politécnica del Litoral
GEI	Gas de Efecto Invernadero
SAT	Sistema de Alerta Temprana
GAR	Gestión Activa de Riesgos
IA	Inteligencia Artificial
RNN	Redes Neuronales Recurrentes
RN	Red Neuronal
PCA	Análisis de Componentes Principales
NN	Redes Feedforward
LSTM	Memoria a Corto Plazo
GRU	Unidad Recurrente Cerrada
BISLTM	Red Bidireccional de Corto Plazo
BIGRU	Red Bidireccional Recurrente Cerrada
RMSE	Error de la Raíz Cuadrada
MSE	Error de la Raíz Cuadrada Media
RR	Coefficiente de Determinación
RFE	Eliminación Recursiva de Funciones
EAR	Eliminación Recursiva de Atributos

SIMBOLOGÍA

m	Metro
Pb	Plomo
Zn	Zinc
µm	Micrómetros
msnm	Metros Sobre el Nivel del Mar
Pa	Pascales
KPa	Kilo-pascales
C	Centígrados
m ³	Metros cúbicos
min	Minuto
% CH ₄	Porcentaje de Metano

ÍNDICE DE FIGURAS

Figura 1.1 Locación de la mina León IV cuenca de Alta Sicilia,Polonia.....	13
Figura. 1.2 Clasificador de una Red Neuronal.	18
Figura 1.3 Estructura de una RNN.....	20
Figura 1.4. Esquema de la predicción de futuros basado en secuencias múltiples interiores por una ventana móvil.	21
Figura 1.5 Arquitectura de la celda de memoria LSTM.....	23
Figura 1.6 Arquitectura de la célula GRU	24
Figura 1.7 Arquitectura De Una Red Bidireccional LSTM.....	25
Figura 1.8 Arquitectura De Una Red Bidireccional GRU.....	26
Figura. 2.1 Desarrollo Del Modelo Predictivo.....	29
Figura 2.2. Ubicación de sensores dentro de la Mina.	32
Figura 2.4 Modelamiento del conjunto de datos en redes neuronales.....	37
Figuras 3.1 Visualización de series temporales y PCA.....	39
Figura 3.2 Función de autocorrelación entre los sensores MM263, MM264, MM256.	41
Figura 3.3 Modelamiento optimizado, función Keras.	42
Figura 3.4 Series temporales: Modelos SRNN, LSTM.....	43
Figura 3.5 Series temporales: Modelos BIGRU, BILSTM.	43
Figura 3.6 Diagrama de dispersión de la predicción de datos del entrenamiento/test en el modelo BILSTM.	44
Figura 3.7 Diagrama de dispersión modelos LSTM, BIGRU.....	45
Figura 3.8 Valores residuales de sensores objetivos en el modelo BILSTM	46
Figura 3.9 Comparación de Modelos Predictivos Recurrentes.	47

ÍNDICE DE TABLAS

Tabla 1.1 Estudios en el desarrollo de SAT vía métodos de aprendizaje profundo... 16	16
Tabla 2.1. Condiciones del método de explotación..... 31	31
Tabla 2.2 Estadística Básica de los datos (Sensores objetivos) 32	32
Tabla 3.1 Relación de atributos y sensores con el Metano..... 40	40
Tabla 3.2 Evaluación del puntaje por Modelo utilizado. 47	47

CAPÍTULO 1

1. INTRODUCCIÓN

En la actualidad, considerando un escenario global con un alto crecimiento demográfico, el consumo energético juega un rol importante en el desarrollo económico y crecimiento tecnológico de las sociedades. Durante años, el carbón mineral ha sido una de las fuentes principales de energía, un pilar en el avance tecnológico para la sociedad. China es el mayor consumidor de carbón mineral, representa aproximadamente el 54.33% de la producción mundial. Dado los efectos negativos del uso de combustibles fósiles, es de esperar que la producción de este mineral disminuya progresivamente; sin embargo, ciertas proyecciones indican que existirá un alza en la producción de carbón para el 2050 en varios escenarios de emisiones. En particular, países en desarrollo tenderán a utilizar este recurso energético debido a que es más económico y viable que otros recursos (Godfrey et al., 2009; Ferreira et al., 2021).

La minería subterránea de carbón es una actividad que puede ser riesgosa, debido a la acumulación de gases que se producen durante la explotación minera. Una gestión activa de riesgos (GAR) aumenta la capacidad para abordar y gestionar riesgos en el futuro; una GAR no espera a que ocurra una lesión o un daño, sino que fomenta la capacitación y entendimiento del problema. Uno de los riesgos más comunes en minería subterránea de carbón es la combustión espontánea en los desechos de carbón, aunque con frecuencia en los materiales de sobrecarga (Joy, 2004; Li y Song, 2009; Onder et al., 2014).

Du et al., (2021), señala que “la minería del carbón tiene como antecedente, estallidos de concentraciones de gas metano; incidentes que tienen un alto índice de víctimas y grandes pérdidas económicas. Desafortunadamente, el proceso de formación que lleva a la combustión del carbón y consecuentemente a la explosión de metano es explicado por fenómenos dinámicos complejos, difíciles de predecir”. En varios sectores industriales, el desarrollo de sistemas de alerta temprana (SAT) ha jugado un rol fundamental en la detección de eventos riesgosos para la seguridad. Según la Oficina de las Naciones Unidas para Reducción de Riesgo de Desastres, el Objetivo de un SAT es poder anticipar los riesgos y amenazas, con el fin de actuar de manera acertada y con

el tiempo suficiente reducir la posibilidad de daños o pérdidas, tanto materiales como humanas. La evaluación consta principalmente de la detección de pronósticos de eventos inminentes, la difusión de información de alerta y la respuesta temprana en situación de riesgo, todo esto basado en una comprensión de la información (Lu et al., 2012).

El desarrollo de nuevas tecnologías de inteligencia artificial y aprendizaje de máquinas ha puesto en evidencia la necesidad de democratizar datos dando cabida al desarrollo de los SAT para el reconocimiento de amenazas potenciales. Los SAT, hasta cierto punto, pueden llevar a cabo una evaluación de riesgos relativamente precisa de los desastres dinámicos de las minas de carbón, pero este tipo de evaluación necesita recopilar datos relevantes manualmente, sin continuidad ni puntualidad. Por lo tanto, es necesaria la instalación de un sistema de monitoreo activo en labores para realizar una evaluación dinámica en tiempo real. Las variables que alimentan a los SAT pueden ser tomadas in situ y son de diversa índole, como datos sísmicos, sonido subterráneo, radiación electromagnética, variables productivas y meteorológicas, entre otras. Cabe recalcar que el SAT es producto de un código y programa que utilizan el aprendizaje profundo para concatenar datos que muestren el comportamiento de algún riesgo con el fin de prevenir y controlar, además de proporcionar soporte en la toma de decisiones (Yuan et al., 2018).

Los métodos del Aprendizaje Profundo se encuentran en un periodo de auge y mejora continua, destacándose en áreas como procesamiento de lenguaje natural y el sistema visual humano, categorizando el dominio del problema por grupos de conocimientos valorados por diferentes datos históricos, aprendiendo los patrones intrínsecos para generar tareas de estimación de manera supervisada o no supervisada. En comparación a sistemas convencionales, el Aprendizaje Profundo brinda respuestas más rápidas que otro tipo de medios, debido a que el aprendizaje del programa es continuo y es de fácil aplicabilidad a sistemas en tiempo real y acceso online (Wang et al., 2022).

Actualmente, el desarrollo de Inteligencia Artificial (IA) sigue mejorando en varios aspectos, desde su estructura y programación hasta sus aplicaciones variadas en diferentes industrias. En la industria minera la IA ha destacado por algunas ventajas que

son aprovechadas por distintas empresas mineras. Podemos señalar que la implementación de IA mediante redes neuronales recurrentes es capaz de interpretar condiciones de tiempo que serán analizadas para el beneficio general. Las principales aplicaciones es en la mejora de equipamiento minero debido a su ayuda en mantenimientos y condiciones físicas de los vehículos que se opera dentro de la mina, también es una gran ventaja la implementación de esta tecnología a los procesos de beneficio siendo capaces de disminuir la utilización de reactivos y aumentar el beneficio mineral así como también es importante en este caso para mejorar la seguridad de las minas de carbón debido a la acumulación de gas metano que es fácilmente detectable y prevenible debido a las lecturas que se establecieron en las mediciones de las mismas (Kang & Choo, 2016).

Considerando que un evento de explosión por metano es un proceso complejo, incluso para ser descrito por modelos físicos y mecánicos, ciertos métodos de aprendizaje profundo tienen el potencial realizar estimaciones con alta precisión utilizando datos de las variables existentes en el ambiente de mina. Este estudio propone el desarrollo de un SAT en una mina subterránea de carbón ubicada en una meseta en Silicia, Polonia, por medio de aprendizaje profundo. Este sistema realiza predicciones secuenciales de concentraciones de metano utilizando variables meteorológicas y de producción. El método a utilizar es un esquema de redes neuronales recurrentes, utilizadas naturalmente para extraer correlaciones temporales de las series de datos (Nguyen et al., 2022).

1.1 Descripción del problema

Históricamente la industria minera subterránea del carbón ha sido asociada a varios accidentes y explosiones debido a la acumulación de gas metano. Las altas concentraciones de metano influyen directamente en la calidad de aire de la mina, así como también en la seguridad y ambiente de trabajo. En este contexto, este proyecto tiene como finalidad analizar y predecir las acumulaciones de este gas mediante el uso de Inteligencia artificial, de cara a mejorar las condiciones de seguridad en la minería subterránea.

1.2 Justificación Del Problema

Este proyecto busca predecir concentraciones de metano en la minería subterránea del carbón, minimizando así los accidentes y pérdidas materiales y de vidas humanas. Se busca mejorar los sistemas de gestión en seguridad minera, de tal forma que, estos puedan ser efectivos y de gran utilidad, garantizando la calidad de aire reduciendo los riesgos y costos de mantenimiento correctivo. Además, se busca mejorar la eficacia del mantenimiento preventivo, específicamente en el proceso de extracción mineral. De igual manera, el modelo de predicción planteado puede ser útil en la automatización de servicios de mantenimiento de maquinaria o en procesos de beneficio mineral con el fin de mejorar la eficiencia de estos utilizando datos previos de gran importancia útiles en la prevención y estudio de accidentes en general.

1.3 Objetivos

1.3.1 Objetivo General

Proponer un sistema de alerta temprana para detección de concentraciones elevadas de metano mediante el uso de aprendizaje profundo aplicado a redes neuronales recurrentes.

1.3.2 Objetivos Específicos

1. Identificar las variables de interés que expliquen el comportamiento de las concentraciones de metano mediante el uso de Aprendizaje profundo.
2. Estimar modelos de predicción para la prevención de explosiones en minas subterráneas de carbón debido a la concentración del gas metano.
3. Diagnosticar el comportamiento del gas metano para su posterior predicción, mediante el uso de redes neuronales recurrentes.

1.4 Marco teórico

1.4.1 Características del sitio de estudio

1.4.1.1 Ubicación de la Mina

La mina subterránea de carbón SZYB LEON IV se encuentra ubicada en la cuenca carbonífera de tierras altas de Silesia, en el Sur de Polonia. Con coordenadas (50.066,

18.438). Esta es una, conocida área mineral donde usualmente se extraen una serie de recursos explotables, como metano, cadmio, plomo, plata, zinc y Carbón. La profundidad del carbón es de aproximadamente 1000 m y contiene alrededor de 70 mil millones de toneladas, con un excelente potencial de extracción. La Cuenca del Carbón de la Alta Silesia incluye el área metropolitana de Silesia como se presenta en la Figura. 1.1 (Ciesielczuk et al., 2021).



Figura 1.1 Locación de la mina León IV cuenca de Alta Sicilia, Polonia.

1.4.1.2 Geología

La meseta carbonífera de Silicia está aproximadamente a 200 - 300 msnm. Esta meseta atribuye su formación al cuaternario con vetas de espesores entre 2 a 7 m de secuencias clásticas del carbonífero superior. La cuenca carbonífera está compuesta con varias decenas de vetas gruesas que a su vez están conformadas por capas de areniscas, lutitas carboníferas y esquistos, siendo esta zona rica en varios recursos minerales como Cd, Pb, Ag, Zn además del Metano.

El inicio de las actividades mineras se remonta al siglo XII, en donde los primeros mineros se beneficiaban del comercio de Ag, Pb y Zn. A finales del siglo XVII y XIX la minería del carbón fue la principal actividad de la zona. Toda esta área ha sido escenario de varios accidentes y explosiones por la acumulación de gas metano (Dubinski et al. 2019). En los sectores norte y oeste de la cuenca carbonífera de Silicia existe una secuencia de rocas pérmicas, así como también calcáreas- dolomitas y calizas. Estas formaciones de rocas y estratos son situados en el carbonífero superior entre el Mioceno y el Cuaternario. En las partes norte y este de la cuenca se identifica la secuencia carbonífera compuesta principalmente por arcillas y lutitas, además de areniscas en menor cantidad. En la parte superior las areniscas forman un estrato delgado cubriendo a la cuenca, las capas del mioceno cubren las formaciones carboníferas con espesores de 90 a 700 m, constituidos por la formación de Skwina compuesta netamente por arcillas con intercalaciones de areniscas (Mendecki et al., 2020).

1.4.2 Descripción de las Operaciones Mineras

a) Método de Explotación: Minería de Tajo Largo (Longwall Mining)

La explotación subterránea de carbón usualmente se realiza explotando grandes cantidades de mineral, que al ser extraídas desprenden material particulado menor a 100µm. El polvo se acumula a lo largo de los túneles o galerías que, usualmente debido a la fricción de la maquinaria con la que se trabaja y el método de explotación (Tajo-largo), generan una excesiva acumulación de este material particulado en las zonas donde se carga y descarga el carbón. El tamaño de las partículas determina usualmente el grado de riesgo que existe en el ambiente de la mina, entre más pequeñas sean estas partículas mayor riesgo existirá de que se provoque una explosión (Eshaghian et al. 2021).

La minería de tajo largo es el método de minería de carbón subterráneo más popular debido a la alta tasa de producción, la menor necesidad de mano de obra y la seguridad del mismo. La extracción de carbón mediante minería de tajo largo es una operación casi continua que implica el uso de soportes de techo hidráulicos de avance automático. La porción excavada está sostenida por unidades de carga: veta de carbón, soporte motorizado. El frente de trabajo en esta minería convencional normalmente adopta una

sección rectangular, de 2–4 m de altura, mediante la excavación de pasadizos en todo su perímetro. En el proceso de extracción, los pilares de barrera de carbón se dejan intactos en ciertas partes de la mina para soportar los estratos superiores. Se permite que el área minada se derrumbe detrás del frente de trabajo, lo que generalmente provoca cierto hundimiento de la superficie (Islavath et al., 2016).

b) Maquinaria Minera

La maquinaria usualmente utilizada en la minería de carbón es el minador de frente largo, esta máquina está dotada de un gran tambor con dientes que se abre paso por la galería extrayendo carbón y depositándolo en una plataforma que lo transfiere a una cinta transportadora que usualmente es acumulado para su extracción, los soportes neumáticos son necesarios para sostener el techo debido al corte del minador. Las excavaciones producen el polvo de carbón altamente explosivo, por ende, se sella la galería con polvo de roca caliza al mismo tiempo que se mina. Entre más largo sea el túnel más complicado es tener la suministración de oxígeno y ventilación necesaria, provocando la acumulación de partículas o gas metano. Luego de operar con el minador de frente largo, la cizalladora de carbón entra en operación; esta maquinaria permite generar una alta producción de carbón de hasta tonelada por segundo, utiliza soportes hidráulicos para protegerse de desprendimiento, además, estos soportes avanzan al mismo tiempo con la cizalladora. Las impurezas que se extraen de la producción de carbón hacen al 30% del total. Finalmente, se lava y seca el carbón para ser utilizado en las diferentes centrales eléctricas para dar energía a los diferentes sectores de la sociedad e industria (Mariño et al., 2015).

1.4.3 Aprendizaje Profundo, Herramientas y Aplicación

1.4.3.1 Inteligencia Artificial

La inteligencia artificial es la capacidad de un programa de aprender y ejecutar acciones al observar una gran cantidad de datos, entrenado su criterio e identificando tendencias para realizar recomendaciones que mejoren las decisiones que tomamos incrementando la eficiencia, productividad y reducción de costos. Algunas ventajas del uso de IA son; repensar procesos, hacerlos adaptables y resolver problemas previamente no resueltos con capacidad de realizar una eficiente labor con máquinas y procesos. Algunas etapas

en minería en los que se puede incorporar IA son: prospección, exploración, beneficio además de maquinaria. La IA se utiliza también en la prevención de riesgos o accidentes debido a que anticipa acciones futuras, mediante el empleo de sensores de seguridad y algoritmos inteligentes la predicción y monitoreo avanzado de las condiciones de trabajo pueden ser de gran ayuda en la prevención de riesgos (Ocaña et al. 2018).

1.4.3.2 Sistemas de Alerta Temprana (SAT)

Los métodos de alerta temprana hasta cierto punto pueden llevar a cabo una evaluación de riesgos relativamente precisa de los desastres dinámicos de las minas de carbón, pero este tipo de método de evaluación necesita recopilar datos relevantes manualmente, sin continuidad ni puntualidad. Por lo tanto, el sistema de monitoreo instalado en la mina para realizar una evaluación dinámica en tiempo real consta de tecnología de toma de datos sísmicos, sonido subterráneo, radiación electromagnética, etc. En sí una serie de datos que nos ayuda a alertar y diseñar un sistema seguro para el ambiente laboral de las minas de carbón (Yuan et al., 2018). En la **Tabla 1.1**, se presentan estudios de desarrollo de SAT bajo diferentes contextos, variando la calidad de aire en interior/ exterior, calidad de agua y procesos químicos industriales. Los métodos de aprendizaje profundo han mostrado versatilidad, brindando estimaciones de un escenario a futuro basado en datos históricos.

Tabla 1.1 Estudios en el desarrollo de SAT vía métodos de aprendizaje profundo

Referencia	Métodos	Campo	Observaciones
(Loy-Benitez et al., 2019)	Redes neuronales recurrentes y con células de memoria (LSTM/GRU)	Calidad de aire subterránea	Se predijeron secuencias de material particulado fino en diferentes horizontes temporales en una estación subterránea.
(Li et al., 2020)	Regresión en procesos Gaussianos	Calidad de aire en exteriores	Se predijo la calidad de aire en dos ciudades utilizando nubes de datos alcanzando más precisión que los benchmarks.
(Na et al., 2018)	Autocodificadores con redes neuronales convolucionales	Procesos químicos en la industria	Se propuso un sistema no lineal subrogado para determinar la probabilidad de muerte luego de un accidente en liberación de gas clorino.
(Asheri Arnon et al., 2019)	Supported vector machines	Calidad de agua	Se desarrolló el SAT para la detección de trazas de contaminantes que pueden resultar mortales.

1.4.3.3 Desarrollo de Modelamiento Predictivo Vía Redes Neuronales Recurrentes

Las redes neuronales recurrentes (RNN) son útiles en varios estudios por su amplia aplicabilidad en campos de investigación. Estas estructuras capturan y procesan información con la ayuda de series temporales, obteniendo así un modelo que procesa información en tiempo real dinamizando los datos por secuencias. Las RNN generan ciclos de procesamiento cuyos resultados son utilizados para analizar el siguiente periodo de tiempo, es así como cada salida es utilizada como entrada del siguiente ciclo y posiblemente de otros ciclos más. Este tipo de red neuronal está diseñado para extraer información contextual definiendo dependencias entre varios ciclos de tiempo, para lograr el proceso de clasificación de datos se establecen estructuras estables que se utilizan de diferentes maneras en secuencias de longitud arbitraria (Loy-Benitez et al., 2020).

1.4.3.4 Redes Neuronales

La red neuronal (RN) es un circuito o red de neuronas utilizadas en la clasificación de datos para resolver problemas de optimización. Una RN requiere de tiempo para su entrenamiento. Estas redes usualmente se emplean para el procesamiento de grandes conjuntos de datos debido a que una RN consta de varias entradas y neuronas ocultas. Los métodos de entrenamiento de redes neuronales se basan en la estimación de peso y sesgo en redes. Cuando un clasificador es grande se requiere de la estimación de varios parámetros con respecto a las muestras de entrenamiento. Cuando tenemos un conjunto de datos grande, generalmente se requiere minimizar la dimensión de los mismos para obtener una red más pequeña que funcione bien con respecto a la precisión de la clasificación de prueba que procesan los datos de entrenamiento. Algunas de las técnicas que se utilizan en la extracción de características influyen directamente en el análisis discriminante y el uso de Análisis de Componentes Principales (PCA) que ayuda a la minimización u optimización (Ec. 1.1) de la función de costo operacional (Kumaraswamy, 2021; Loy-Benitez et al., 2019).

Ec. 1.1

$$X = \sum_{x=1}^m \lambda_x = \frac{1}{2} \sum_{x=1}^m \sum_{k=1}^n (P_{xk} - A_{xk})^2$$

Donde “x” representa al número patrón, “m” representa al tamaño de la muestra, P_{xk} es una salida deseada, A_{xik} es la producción real de neuronas y “n” representa la neurona de salida. En general, este método de optimización es el más utilizado en el clasificador de RN. La Figura 1.2 presenta la arquitectura de una RN (Lasso & Kazanzides, 2019).

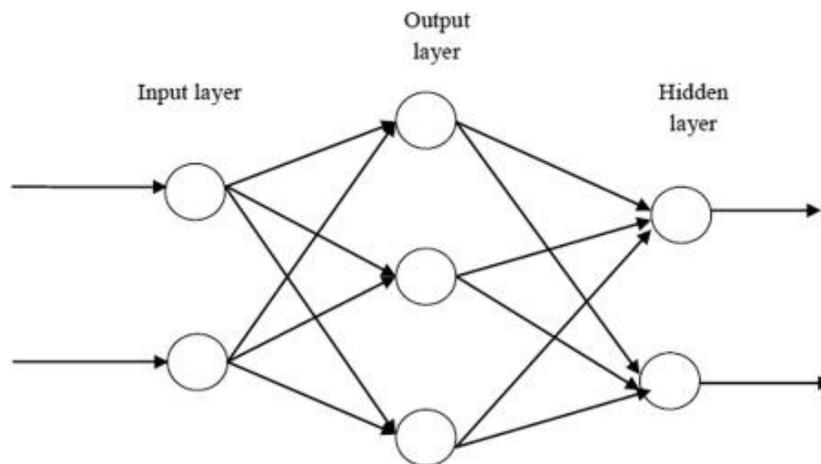


Figura. 1.2 Clasificador de una Red Neuronal.

En general, podemos interpretar una RN como un clasificador que consta de una serie de procesadores que funcionan en paralelo y están dispuestos en forma de niveles en donde nodos conforman parte de la red, los cuales están fuertemente relacionados e interconectados entre sí al usar los nodos del nivel anterior con los del siguiente nivel. Sin embargo, cada nodo tiene su propio funcionamiento que influye directamente en la programación y aprendizaje por sí mismo. Las RN son una herramienta atractiva en la clasificación de datos (Sharma et al., 2022).

1.4.3.5 Redes Neuronales Recurrentes Y Series Temporales

Las redes neuronales recurrentes (RNN) son RN que se usan ampliamente en el proceso de análisis de secuencias debido a que esta red está diseñada para extraer información de tipo contextual que define dependencias entre varias marcas de tiempo. La RNN consta de varias capas recurrentes sucesivas que modelan secuencialmente distintas capas que sirven como modelo para mapear otras secuencias, tienen una capacidad alta de capturar datos en secuencias, estas señales contextuales son estables en la secuencia de la red y se utilizan de manera efectiva para lograr un proceso de clasificación de datos cabe recalcar que las RNN operan en secuencias con longitud

arbitraria. Estas redes son una extensión de las Redes Feedforward NN con la presencia de bucles en capas ocultas en donde la entrada de datos es una secuencia de muestras que identifica la relación de tiempo entre las muestras. El nodo RNN consta de un único sesgo y peso, se evalúa usando la unidad recurrente cerrada presente en la Figura 1.3. La configuración de la red se forma utilizando los parámetros de red en donde cada dato de entrada genera la salida en una secuencia de tiempo específico. La optimización de los parámetros de la red se logra empleando el radiante estocástico con base en la secuencia de datos de entrada. Sin embargo, los hiperparámetros son la estructura de la red, es decir: tamaño, capas, la longitud de la secuencia, el tamaño del conjunto de datos, la tasa de aprendizaje (Kumaraswamy, 2021).

La capa recurrente consta de la función recurrente, que toma el vector de entrada Y_x y la unidad oculta del estado anterior Z_x como entrada y genera un nuevo estado oculto como:

Ec. 1.2

$$z_x = d(y_x, Z_{x-1}) = \tanh (P * y_x + Q * z_{x-1})$$

Unidades de salida:

Ec. 1.3

$$V_x = \text{soft max}(R * Z_x)$$

Donde P y Q representan las matrices de peso, "Y" la función de activación tangente hiperbólica. RNN usa la función altamente complicada para controlar y aprender el flujo de información en la capa recurrente capturando dependencias a largo plazo.

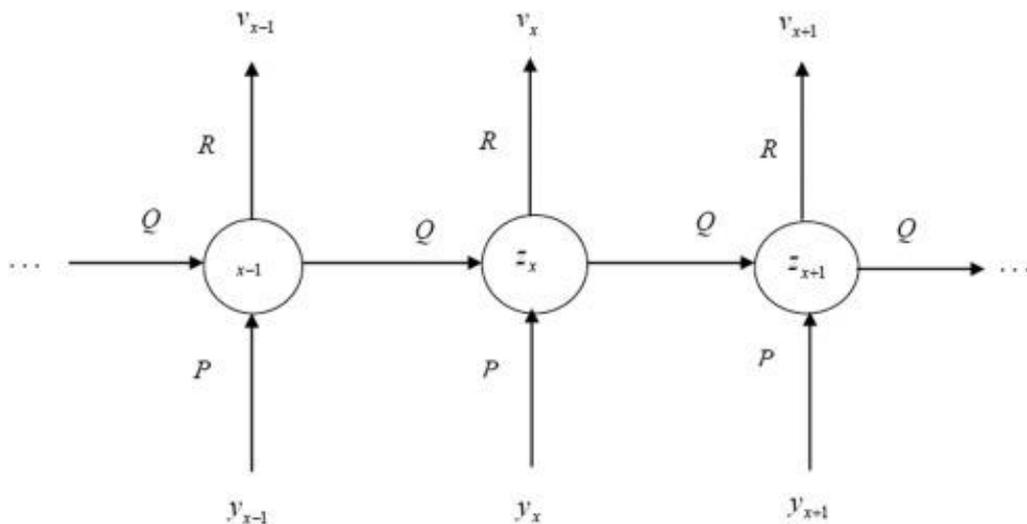


Figura 1.3 Estructura de una RNN.

1.4.3.6 Desarrollo de Ventana Móvil.

La realización de un aprendizaje secuencial tiene un alto impacto en este estudio, sumado a la predicción de un nuevo valor basado en los anteriores. Las variables son como la ya conocemos el conjunto de datos brindado por los sensores. Para iniciar con el proceso predictivo se debe definir varias secuencias de información pasada a partir del conjunto de datos de la entrada original, puntualizando "K" puntos de datos retrospectivos, lo cual representa la entrada del modelo secuencial. La salida es definida como un número de conjuntos de datos "L" futuros que se predecirán. En la Figura 1.4 "X" representa la salida en una secuencia de longitud "L". El cálculo del valor futuro proviene de un análisis en el retraso que se basa o tiene raíz en la función de auto correlación (Y_k) basada en el conjunto de datos pudiendo determinar el dato que tiene correlación asociando esto a la longitud de ventana con los puntos "K" (Loy-Benitez et al., 2020).

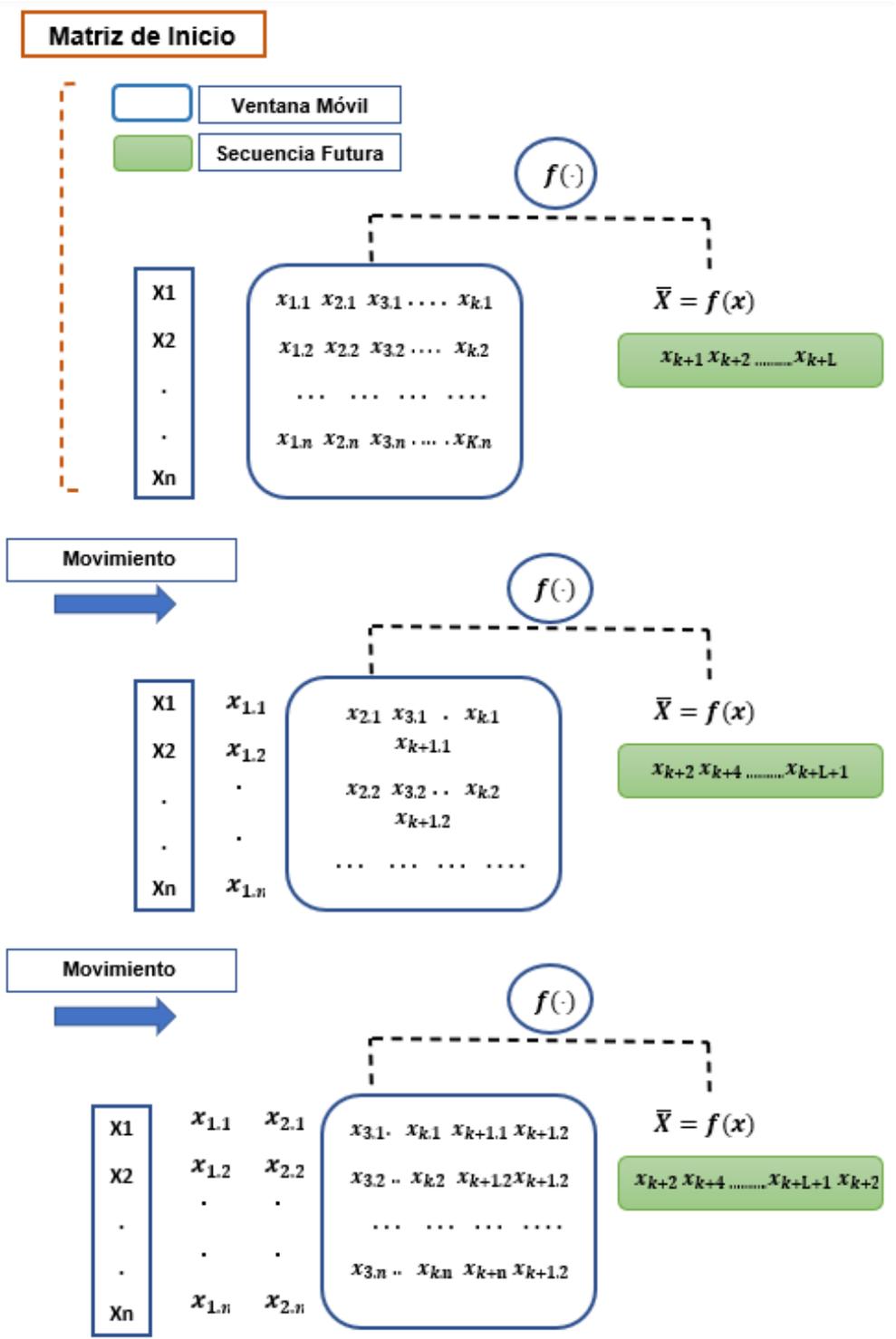


Figura 1.4. Esquema de la predicción de futuros basado en secuencias múltiples interiores por una ventana móvil.

1.4.3.7 Unidad Recurrente Cerrada (GRU) / Memoria a Corto Plazo del Modelo Predictivo y Modelos Bidireccionales.

a) Long short term memory (LSTM)

Long short term memory (LSTM) es un tipo de red neuronal utilizada principalmente para el modelamiento de series temporales con una dependencia a largo plazo. Son capaces de controlar la propagación interna de la red incluyendo celdas y puertas de memoria. LSTM son útiles debido a su capacidad de retener o descartar información, esto mejora significativamente la precisión de predicción con dependencias a largo plazo. Consta de puertas principales denominadas: puerta de entrada (In_t), puerta de salida (Out_t) y puerta del olvido (Fo_t), además se muestra el cálculo del estado actual de la celda ($Cell_t$) y la salida del estado oculto (h_t) (Loy-Benitez et al., 2020), las cuales se describen a continuación:

Ec. 1.4

$$In_t = \sum(w_i * [hid_{t-1}, x_{t-1}] + b_i)$$

Ec. 1.5

$$Out_t = \sum(w_o * [hid_{t-1}, x_{t-1}] + b_o)$$

Ec. 1.6

$$Fo_t = \sum(w_f * [hid_{t-1}, x_{t-1}] + b_f)$$

Ec. 1.7

$$Cell_t = F_t * Cell_{t-1} + In_t * \tanh(w_c * [hid_{t-1}, x_{t-1}] + b_c)$$

Ec. 1.8

$$h_t = Out_t * \tanh(Cell_t)$$

Donde:

w_i, w_o, w_f, w_c son variables que representan las matrices de peso. Los vectores de sesgo se representan como b_i, b_o, b_f, b_c ; hid_{t-1}, x_{t-1} que representan a la celda LSTM con su entrada y salida, respectivamente para un momento t-1 todo en conjunto con la función de activación *sigmoid*. La arquitectura de esta celda LSTM se muestra en la Figura 1.5.

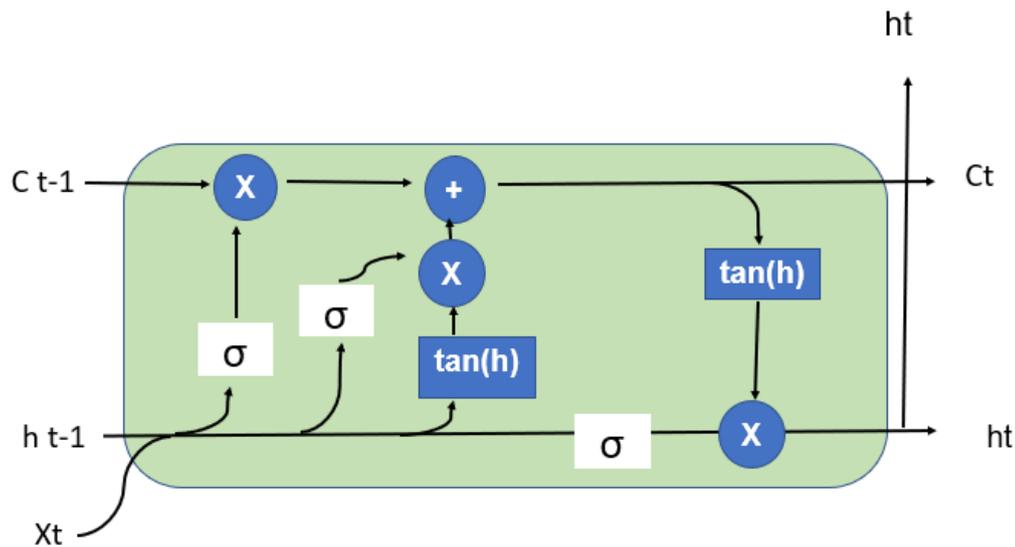


Figura 1.5 Arquitectura de la celda de memoria LSTM

Y su funcionamiento corresponde a:

1. C_{t-1} Especifica una entrada nueva para el cálculo de estado actual de la celda ($Cell_t$).
2. σ Establece el valor de la memoria anterior que debe eliminarse del estado de celda.
3. Se evalúa la representación oculta de la celda mediante $\tan(h)$.
4. Se visualiza la salida considerando $h_t = h_{t-1} * c_{t-1}$; tomando a X como una matriz de proyección utilizada para reducir la dimensión de la salida del estado oculto de la celda.

b) GRU

Una unidad recurrente cerrada es una mejora de lo que conocemos como LSTM. Esta, incluye algunas ventajas como una puerta de actualización y así como también de reinicio. GRU fusiona la ventana de olvido de LSTM con la puerta de entrada creando la puerta de actualización. Otra ventaja destacable es la combinación del estado oculto y de celda, esto implica la posible solución al problema del gradiente de fuga. GRU plantea controlar la información de la unidad y transferirla a otra red, así como también un procesamiento rápido con conexión entre datos secuenciales ilustrada en la (Figura. 1.6).

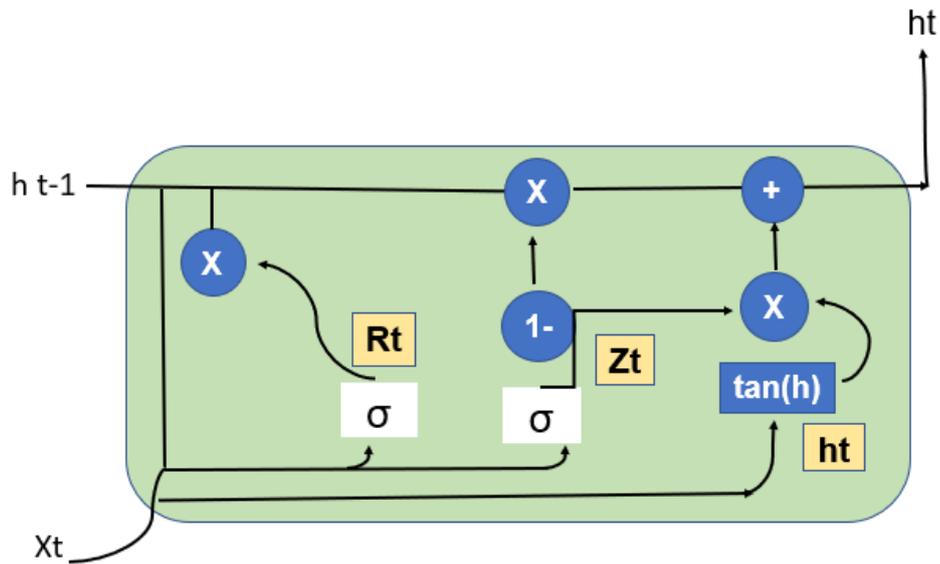


Figura 1.6 Arquitectura de la célula GRU

Las puertas internas son calculables como:

Ec. 1.9

$$Z_t = \sum ([hid_{t-1}, x_t] + b_z)$$

Ec. 1.10

$$R_t = \sum ([hid_{t-1}, x_t] + b_r)$$

Ec. 1.11

$$h_t = \sum ([hid_{t-1}, x_t] + b_r)$$

Ec. 1.12

$$h_t = \tanh [X * hid_{t-1} + R_t * (X * hid_{t-1}) + x_t]$$

Ec. 1.13

$$hid_t = (1 - z_t) * hid_{t-1} + (z_t * h_t)$$

X, hid_{t-1} representan las matrices de peso; R_t, Z_t, h_t son los vectores de sesgo; σ controla la información del estado actual; h_t el vector candidato y $z_t * h_t$ la operación del producto, respectivamente (Loy-Benitez et al., 2020; Nguyen et al., 2022).

c) Modelos Bidireccionales LSTM y GRU

En la Figura 1.7 se muestra la arquitectura de una red bidireccional. Se observan dos bloques de contexto, uno con recurrencia hacia el pasado y otro con recurrencia enfocada en el futuro en cada tiempo t . La transformación hacia la izquierda del modelo pasa la información al tiempo anterior y la derecha al tiempo posterior la transformación del medio es para información del tiempo actual. Esto hace que se puedan analizar problemas donde exista relación entre partes de la secuencia, relaciones que nos tienen que ser iguales Keras provee una celda que ya combina las dos redes recurrentes (la que procesa la primera orden y la que procesa en orden inverso) y que integra en una sola salida ambas respuestas, la forma de crear una RNN (LSTM en este caso) bidireccional en Keras se detalla a continuación. Donde “ x ” es un vector y representa la entrada, “ s ” es vector y representa la capa oculta, “ U ” es la matriz de peso, “ O ” es una representación de salida, V es la matriz de peso de la capa oculta (Nguyen et al., 2022).

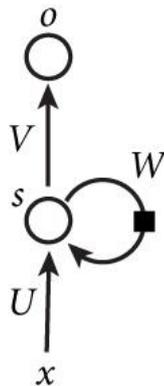


Figura 1.7 Arquitectura De Una Red Bidireccional LSTM

La red neuronal BiGRU es empleada para almacenar y recuperar información durante largos periodos de tiempo como observamos en la Figura 1.8, su desempeño en la práctica puede ser no tan satisfactorio como en la teoría, debido a que la red solo accede a información pasada. Esta red bidireccional tiene una capa futura en la que la secuencia de datos se encuentra en la misma dirección, pero con opuestos sentidos. Así la red utiliza capas ocultas para extraer información del pasado como del futuro conectándose entre sí en una estructura bidireccional. Por consecuencia ayuda a las redes neuronales a extraer más información y mejorar el rendimiento del proceso de aprendizaje (Deng et al., 2019).

Donde “X” son nuestras variables de entrada y “O” las de salida recorriendo y recolectando información del futuro y pasado.

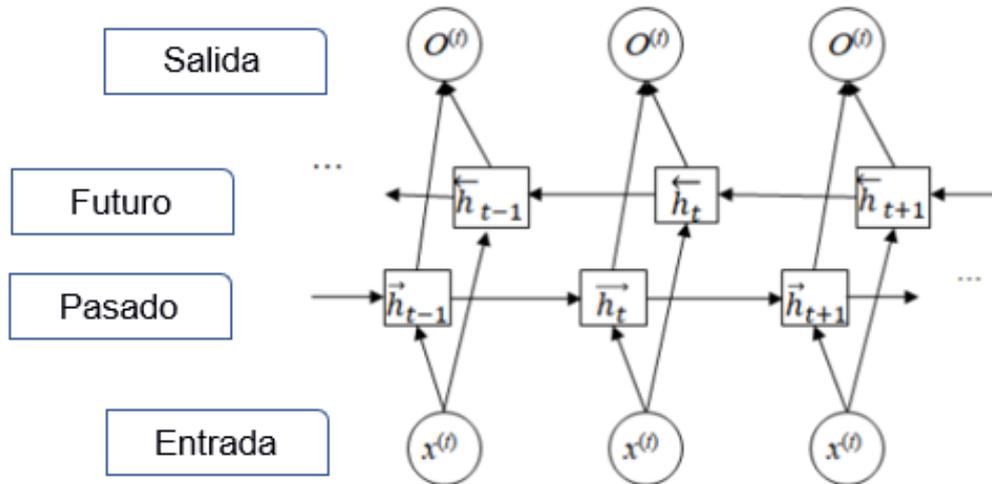


Figura 1.8 Arquitectura De Una Red Bidireccional GRU

1.4.4 Evaluación de modelos predictivos

a) Error de la Raíz Cuadrada Media (RMSE).

Una vez completado el modelo predictivo es necesario evaluarlo mediante la calidad y tipo de modelo, así como su capacidad de procesamiento. Las métricas son necesarias para evaluar y reemplazar el mejor modelo que proyecte la situación a futuro de algún evento en específico. Este criterio se puede utilizar para medir el porcentaje de variabilidad total de la respuesta explicada por el modelo además es necesario para medir el porcentaje de variabilidad total de la respuesta explicada por el mismo (Pougnnet et al. 2015; Loy-Benitez et al. 2020).

Ec. 1.14

$$RMSE = \sqrt{\frac{1}{m} * \sum_{i=1}^m (X_i - Z_i)^2}$$

Donde X_i son datos observados y analizados por la Inteligencia Artificial.

b) Error Cuadrático Medio (MSE).

Corresponde al error de predicción por cuadrado cuyo criterio es:

Ec. 1.15

$$MSE = \frac{1}{m} \sum_{i=1}^m (Y_i - \bar{Y}_i)^2$$

Donde m representa el número de elementos de datos en el proceso de prueba. El criterio mide el error cuadrático medio que se analiza en el proceso de prueba, estos términos prácticos, un valor bajo del criterio MSE indica que los valores predichos se encuentran cerca a los reales. Ocasionalmente se relaciona el RMSE y MSE en la siguiente ecuación (Ait-Amir et al., 2015).

Ec. 1.16

$$RMSE = \sqrt{MSE}$$

c) Coeficiente De Determinación.

Un diagnóstico inicial del modelo se realiza analizando la validez de este a través de los valores del modelo asociado a los puntos de información crítica del proyecto. Restringiendo el número de simulaciones es posible definir una colección de puntos de prueba necesarios para la evaluación del error (Ait-Amir et al., 2015). Esto se visualiza matemáticamente de la siguiente manera:

Ec. 1.17

$$R^2 = 1 - \frac{\sum_{i=1}^m (X_i - Z_i)^2}{\sum_{i=1}^m (X_i - \dot{X})^2}$$

“Yzi” son los datos reconstruidos y Xi son datos observados y analizados por la Inteligencia Artificial (Nguyen et al., 2022).

1.4.5 Función De Auto correlación

La función de auto correlación es un valor calculado por un código recurrente, este se usa para representar que tan similar es un valor dentro de una serie de tiempo a un valor anterior. La auto correlación, en términos simples, es una señal consigo misma en diferentes retrasos definiendo esto como:

Ec. 1.18

$$R_{ff}(l) = \sum_{n=0}^N f(n)f(n-1)$$

Donde “n” es cada son los intervalos de secuencia. El Máximo retraso es 0, si los inter-valores de auto correlación son pares podemos suponer que existe un componente periódico en ese intervalo. Utilizamos esta función para identificar tendencias en datos y la influencia de los valores observados anteriormente en una observación actual, es muy útil para pronosticar valores futuros de los datos históricos como con los que trabajamos (Loy-Benitez et al., 2022).

A continuación, se detalla el proceso que se cumple en la utilización de esta función:

1. La generación de datos de series temporales, útil para la exploración de herramientas de análisis como modelos estadísticos; datos que deben tener una secuencia entre sí.
2. Verificación del conjunto de datos mediante valores booleanos que indican Falso (0) si existe datos faltantes o verdadero (1) para aquellas secuencias que estén completas.
3. Reemplazo de datos faltantes, en sí un método de interpolación que corrige los datos que faltan con valores dentro de un rango generalmente cercano.

CAPÍTULO 2

2. METODOLOGÍA

A continuación, se presenta una estructura para el desarrollo del estudio, el cual se encuentra dividido en 3 fases críticas del modelo predictivo. La Figura. 2.1 muestra el proceso para llegar a la selección del método exitoso.

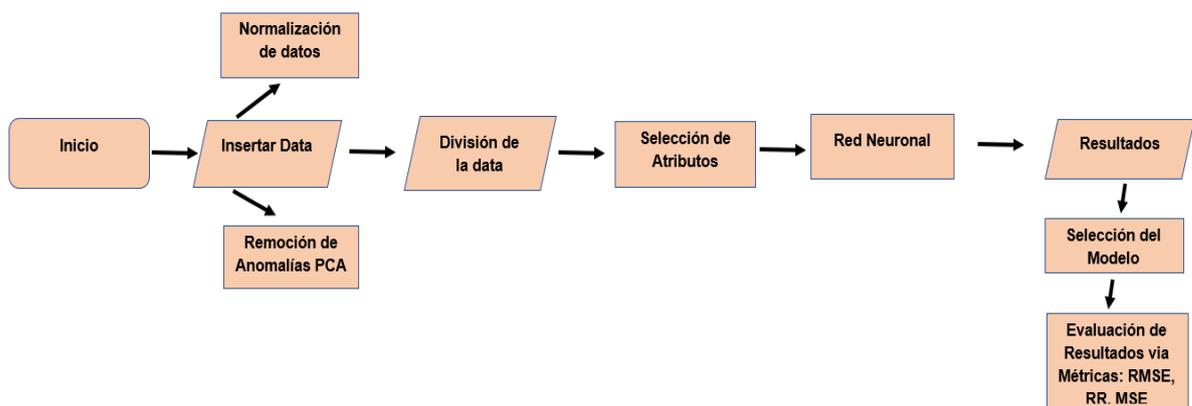


Figura. 2.1 Desarrollo Del Modelo Predictivo

a) Fase 1

El pretratamiento de la base de datos busca eliminar todos los elementos erráticos, así como también normalizarlos para una mayor eficacia en los procesos neuronales predictivos. La selección de atributos se usa para medir o evaluar los atributos más importantes al modelo predictivo, eliminando atributos redundantes o sin relevancia en el modelo predictivo.

b) Fase 2

Entrenamiento de redes neuronales para el modelo predictivo de la concentración de metano evaluado en tres sensores de la mina de carbón, evaluando el rendimiento del modelo predictivo.

c) Fase 3

Esta última fase busca relacionar los mejores resultados predichos por los modelos, evaluarlos mediante las tres distintas métricas y comparar dicha evaluación en cada modelo para decidir cuál es el que mejor eficiencia tiene, es decir, el que mejor predicción genera.

2.1 Descripción de los datos

La data inicial fue obtenida por 23 sensores que proporcionaron datos por cada segundo analizado, sumando más de 9 millones de datos, relacionando así 27 variables que comprenden condiciones internas y externas de la mina, así como también lecturas sobre los cambios en concentración de gases que comprenden las mediciones de las condiciones climáticas, además del monitoreo de las condiciones de control de gases obtenidos por los sensores cuyos valores son relacionados a la temperatura, presión y velocidad de gases. El tamaño de este conjunto de datos ocupa un significativo espacio de almacenamiento y procesamiento para nuestra red, sin embargo, los datos recopilados son útiles en el estudio del diagnóstico de una red neuronal en tiempo real debido a su naturaleza multidimensional. Es importante detallar el estudio de sensores objetivos, los cuales están presentes en la Tabla 2.1 y 2.2 (MM263, MM264, MM256). Estos sensores brindan más información sobre los niveles de metano con la finalidad de alertarnos de un peligro inminente en las concentraciones de este gas y valores altos de concentración.

Para maximizar la eficiencia de la red se realizó un promedio de datos convirtiéndolos en conjuntos de una hora. La aplicación de una estandarización de datos es importante debido a que los valores obtenidos por los sensores en muchos casos se encuentran en diferentes métricas y rangos haciendo imposible el procesamiento de estos, es por ello que aplicamos la estandarización para una tener una correcta precisión del sistema de alerta temprana. Además, el proceso de estandarización permite eliminar anomalías y datos erráticos que afectaban directamente en la precisión de nuestra red neuronal. La Tabla 2.1 detalla 3 sensores que brindan información importante de datos útiles para nuestro procesamiento debido a que estos son en sí una métrica para el control de gas metano en las galerías. Posteriormente se analiza los datos haciendo un entrenamiento de una RN, básicamente podemos observar el rango en el cual los sensores han analizado el ambiente receptaron datos y de proceso. Entender la función que cada uno de estos sensores realiza, Además de, sus medidas y atributos indican qué tipo de clasificación tiene el sensor junto a las unidades de los valores que refleja cada uno de estos datos.

Tabla 2.1. Condiciones del método de explotación.

Información de Sensores (Colector de información y Atributos)				
Sensor	Denominación	Tipo	Rango de Medición	Unidad
MM252	Métrica de metano	Apagado	-0.1 / 30	[%CH4]
MM261			0 / 30	
MM262			-0.2 / 30	
MM263*			-2 / 30	
MM264*			-2 / 40	
MM256*			0 / 30	
MM211			-2 / 30	
CM861	Altas concentraciones de CH ₄	Registro	-0.2 / 67.7	[%CH4]
CR863			-8 / 258	[Pa]
7 P_864	Presión dentro del drenaje de metano CH ₄ .		0 / 435.4	[kPa]
TC862	Temperatura dentro del drenaje		0 / 40.5	[C]
WM868	CH ₄ according to CM, CR, P, TC		0 / 6.39	[m3/mi]

(*) Sensores objetivos; estos sensores son de gran utilidad para el modelo predictivo, muestra los niveles de metano por cada segundo dentro de la mina.

La Tabla 2.2 explica y visualiza las medidas de dispersión estadística de los sensores que principalmente intervienen en el modelo predictivo, los datos recolectados por estos sensores son %CH₄ y servirán como métrica de concentración. La media de los sensores objetivos tiene valores que oscilan en diferentes rangos, lo cual significa que esta variable es afectada por los valores erráticos y grandes. La varianza de estos nos indica que estos no varían mucho entre sí.

Tabla 2.2 Estadística Básica de los datos (Sensores objetivos)

Medidas de dispersión (Conjunto de Datos)				
Sensor	MM263	MM264	MM256	Unidades
Media	0.2547	0.3274	0.4297	[%CH4]
Moda	0.2000	0.2000	0.2000	
Mediana	0.2175	0.3000	0.4018	
Rango	(0-1.0564)	(0-1.1192)	(0-1.2696)	
Varianza	0.0249	0.0294	0.0305	
Desviación Estándar	0.1577	0.1714	0.1746	

Los datos de las condiciones del método de explotación que reflejan los sensores objetivos están en un área en la cual se acumula carbón residual, que, sin tomar las medidas necesarias, pueden generar una gran concentración de CH₄. Las concentraciones pueden provocar fracturas y largos periodos de difusión, afectando la seguridad del personal de la mina. Gráficamente se representa la ubicación de los sensores en el siguiente croquis con la numeración 1 del Sensor Map (Figura. 2.2).

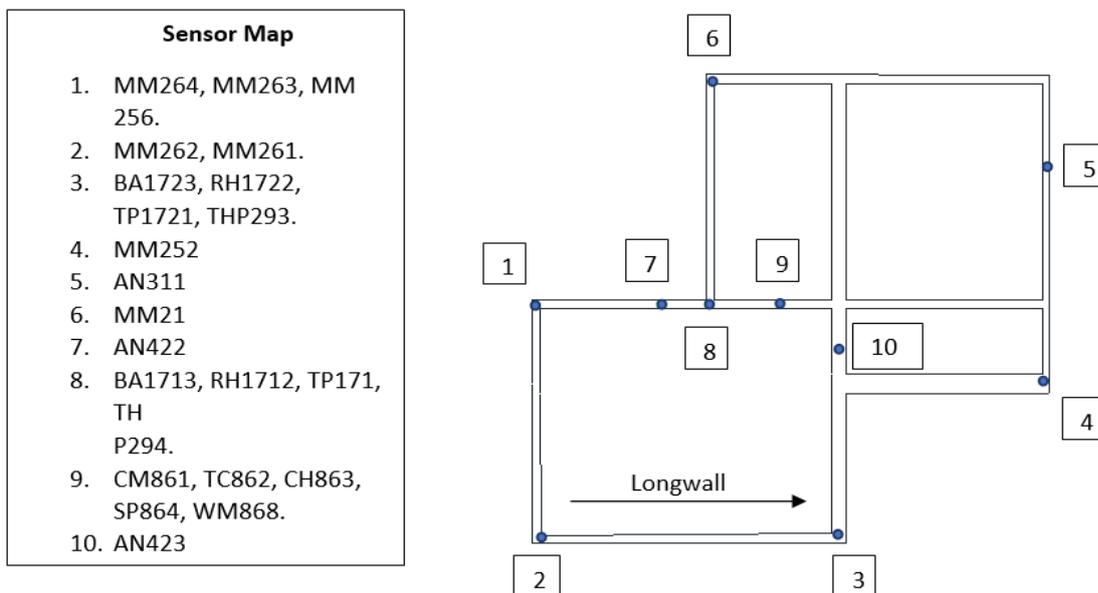


Figura 2.2. Ubicación de sensores dentro de la Mina.

2.2 Sensores y Características.

Los sensores de gas metano que se utilizaron son diseñados para realizar una detección de gas, adecuando su sensibilidad para una lectura segura. El sensor tiene una salida digital por un comprobador con un umbral ajustable para trabajar con cualquier microcontrolador. Al recolectar información estos sensores captaron valores de 6 atributos cada uno, estos son: año, mes, día, hora, minuto, segundo. Los datos son recopilados y presentados como un archivo csv. Los sensores que activan los umbrales de advertencia en el modelo de predicción son: MM263, MM264, MM256; los datos captados por estos sensores representan a la cantidad de metano presente en el ambiente de trabajo. El frente de operaciones en una mina de carbón es el lugar con mayor material particulado debido al roce de la maquinaria utilizada en el proceso de extracción mineral (Neira, 2015). Kozielski et al., (2021), señala que “las características principales de estos sensores corresponden a su manejo con una elevada sensibilidad del gas metano en un amplio rango, además, detecta propano e hidrogeno. La larga duración y el bajo costo es importante debido al manejo simple del circuito lo que lo hace un sistema ideal y activo para la valoración y estudio de datos necesarios para crear el modelo de predicción planteado”.

2.3 Herramientas y Programa Usado.

Se inició creando un entorno de desarrollo de programación multiplataforma de código abierto, para programación científica en Python. Este programa llamado Spyder posee funciones de edición y pruebas interactivas, con un desarrollo informático numérico. Se usó el entorno Anaconda Navigator (anaconda3) para programar con Spyder, en el código en sí, se inició llamando librerías (Pandas, Matplotlib, Sklearn, Numpy, Seaborn) necesarias para la lectura de nuestro conjunto de datos; se planteó un arreglo de estos datos escalando por rangos óptimos necesarios para el análisis de concentraciones altas de metano. Se correlacionaron las variables a través del método PCA y se hizo una eliminación de anomalías para el posterior entrenamiento de modelo predictivo, consiguiente se realizó pruebas al modelo. El resultado de la predicción es un diagrama que fija la comparación de los niveles medidos y la proyección de nuestro modelo predictivo para corroborar la información y observar que tan preciso y eficaz es el modelo (Damian, 2015).

2.4 Detección de Anomalías Vía análisis de Componentes Principales.

En la recolección de datos de la mina lo más frecuente es tomar el mayor número posible de variables. Sin embargo, si tomamos demasiadas variables estos datos serán multivariados y de diversas distribuciones, afectando la eficacia del modelo predictivo. Asumiendo que la reducción de dimensión de los datos se ha realizado de forma que la mayoría de los datos (normales) queden bien representados, aquellas observaciones con mayor error de reconstrucción deberían ser las más atípicas y fácil de reconocer para su eliminación del análisis. Utilizando el PCA (Ec. 2.1; 2.2; 2.3) eliminamos aquellos datos que no tengan relación entre otros. La detección de anomalías es la técnica para identificar eventos raros u observaciones que se pueden generar al ser estadísticamente diferentes del resto de las observaciones en los datos. Tal comportamiento anómalo generalmente se traduce en algún tipo de problema (Kotarba et al., 2021).

Ec. 2.1

$$Y = AX$$

$$Y = \begin{pmatrix} y1 \\ y2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ yp \end{pmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1P} \\ a_{21} & a_{22} & \cdot & \cdot & a_{2P} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{p1} & a_{p2} & \cdot & \cdot & a_{pP} \end{bmatrix} \quad X = \begin{pmatrix} x1 \\ x2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ xp \end{pmatrix}$$

La matriz de covarianza es:

Ec. 2.2

$$\text{Var}(y1) = \lambda1, \text{Var}(y2) = \lambda2,$$

$$\text{Var}(yp) = \lambda p$$

Ec. 2.3

$$\text{Donde} \quad A = \begin{bmatrix} \lambda1 & 0 & 0 \\ 0 & \lambda1 & \cdot \\ \cdot & \cdot & \lambda p \end{bmatrix}$$

El método de PCA es simplemente la suma de los cuadrados de estos K-términos, por lo que una gráfica de contribución residual, más convenientemente mostrada como un gráfico de barras de estos K-términos que indican cuál de las variables originales están

más asociadas con la desviación fuera del plano del modelo (Jason, 2018; Kotarba et al., 2021).

2.5 Selección de atributos vía eliminación recursiva de funciones (RFE)

Algunas variables aportan información repetida sobre el problema debido a una alta correlación entre sí. Este estudio emplea el método RFE, que consiste en un algoritmo de selección de características puntuando su nivel de relevancia, proporcionar una forma de calcular puntajes importantes, como un árbol de decisión. El algoritmo utilizado en RFE es un método que tiene un alta importancia, ya que permite realizar la selección de características para problemas de modelado predictivo de clasificación y regresión; se pueden utilizar diferentes algoritmos y en general estos ayudan a reducir el costo computacional (Jason, 2018).

2.6 Procesamiento de Ventana Móvil y su Determinación Vía función de Auto correlación.

Para iniciar con el proceso predictivo se debe definir varias secuencias de información pasada a partir del conjunto de datos de la entrada original, puntualizando "P" puntos de datos retrospectivos, lo cual representa la entrada del modelo secuencial. La salida es definida como un número de conjuntos de datos "L" futuros que se predecirán. El modelo bidireccional funciona mediante el procesamiento de datos antiguos y datos futuros haciendo que el modelo pueda tener una mayor presión en su predicción. La auto correlación observa dos variables x y y , se analiza x_t versus x_{t-l} , donde, l son los "retardos", con $l = 1$ se estaría comparando cada observación con su inmediata anterior, con $l = 2$ se estaría comparando cada observación con la que antecede a su inmediata anterior y así sucesivamente (Bonet Cruz et al. 2007).

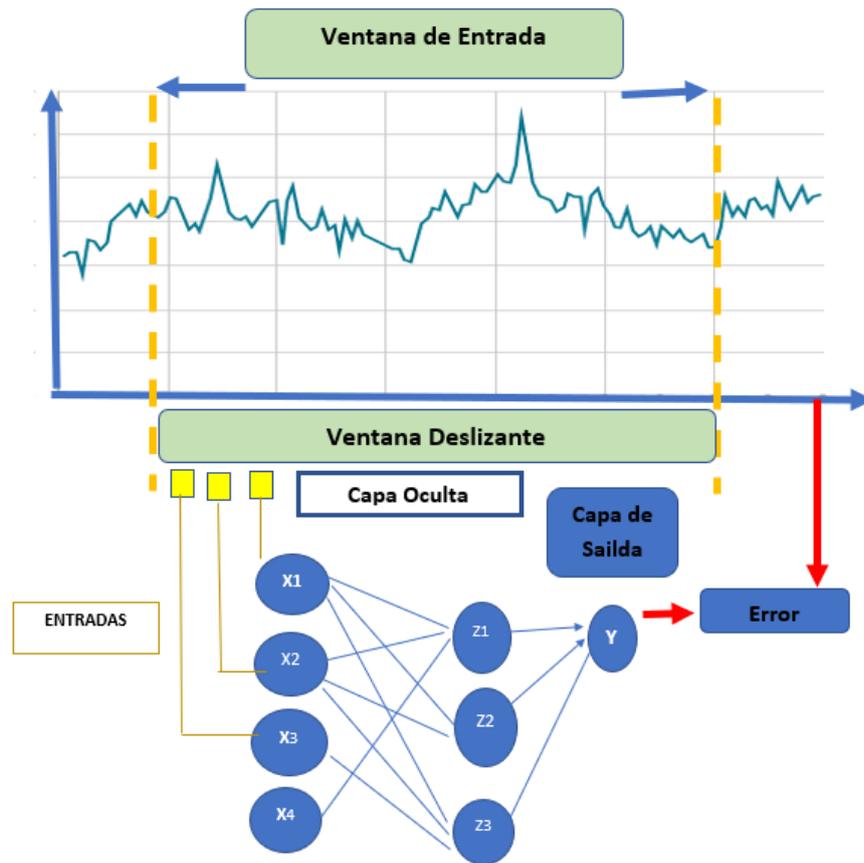


Figura 2.3 ventana móvil

2.7 Arquitectura de la Red Neuronal

Las entradas utilizadas, corresponden a los datos recolectados por los sensores objetivos cuya finalidad es brindar la información necesaria para cada proceso dentro del modelo predictivo. Con el fin de mantener una memoria a corto plazo para diferentes operaciones el modelo puede conservar resultados para posteriores operaciones dentro del modelo predictivo. Para este modelo, se va a ajustar "KerasTuner". Este parámetro brinda información de la cantidad de filtros y convoluciones a utilizar en un muestro del tipo ventana rodante en donde, se tendrá siempre cortes y epochs; para cada muestra se mantiene estos valores con la finalidad de calcular la efectividad del modelo escalando la data para escoger la combinación a usar sobre la data total.

2.8 Desarrollo de Sistema de Alerta Temprana Vía Redes Neuronales Recurrentes

Agrupando los procesos planteados en este proyecto se pudo idear las bases del sistema de alerta que puedan generar avisos de manera segura sobre las condiciones en las que los procesos mineros se llevan a cabo dentro de una mina de carbón y cuáles no son posibles realizar debido a los altos porcentajes de concentración de metano. Evaluadas las métricas de rendimiento del modelamiento predictivo como MSE, RR Y RMSE, la eficiencia de los modelos es contrastada en una evaluación de advertencia de riesgo. Los niveles de riesgo para las operaciones mineras y los trabajadores que se encuentren dentro de la mina, las cuales son una medida de fidelidad para observar cuán acertado es el valor de predicción con respecto a la variable real. El modelo general de obtención y procesamiento de datos y su modelación es detallado en la Figura 2.4.

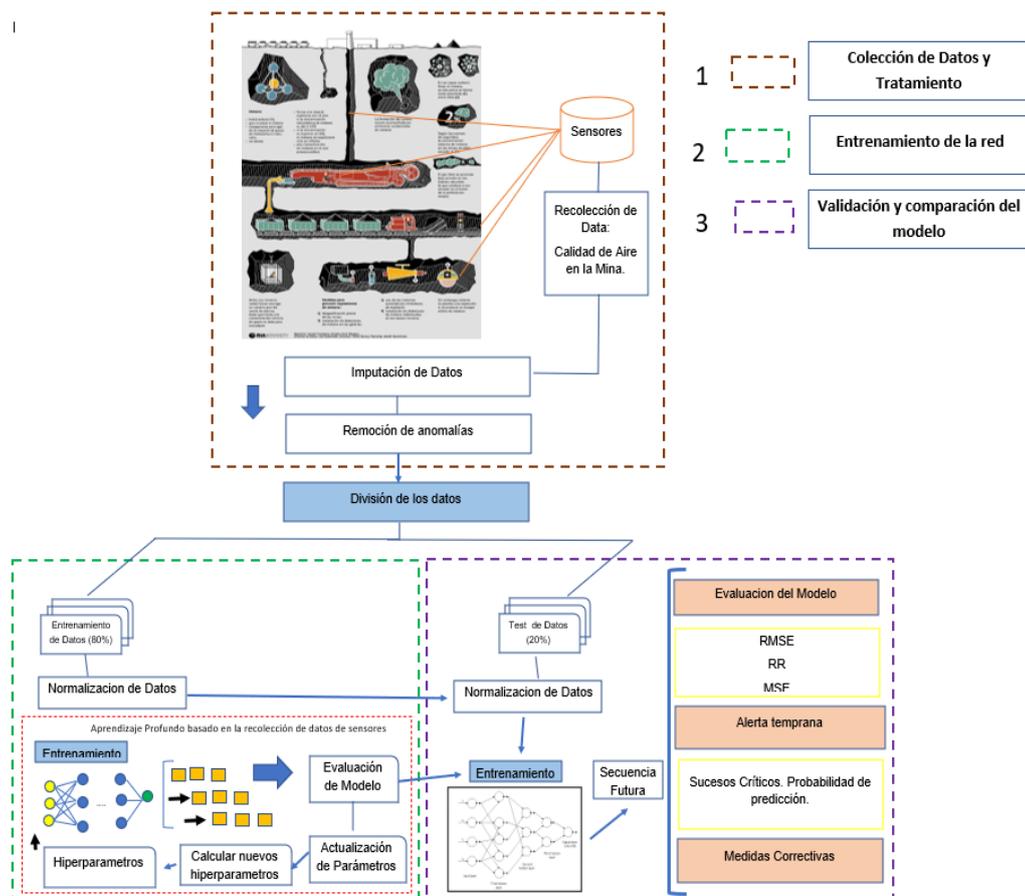


Figura 2.4 Modelamiento del conjunto de datos en redes neuronales.

CAPÍTULO 3

3. RESULTADOS

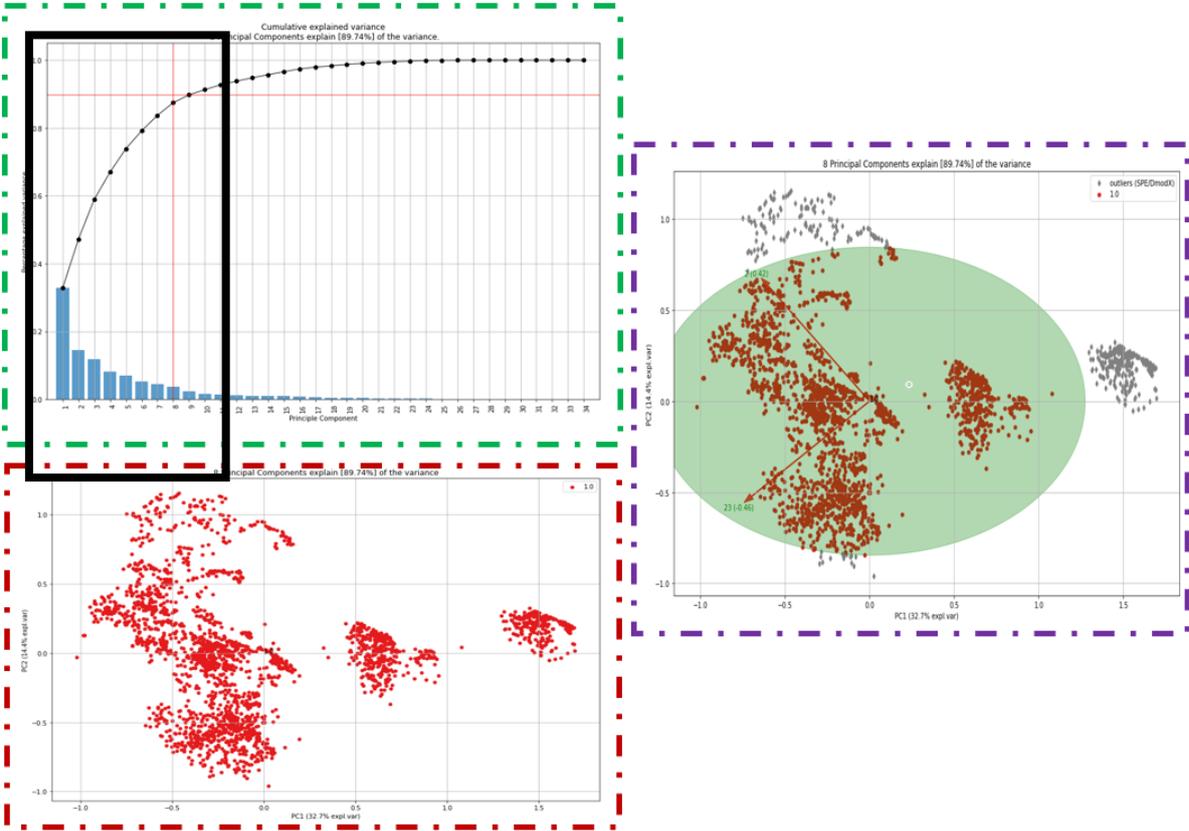
El modelo propuesto se aplicó a la predicción de concentraciones de metano en una mina subterránea de carbón localizada en Polonia. Se trabajó con 9 millones de datos aproximadamente, de este conjunto un 80% fue utilizado para el entrenamiento de la RN, mientras que el 20% se empleó como Test para evaluar el comportamiento de la RN con datos no antes vistos. El primer proceso fue la normalización de datos que tuvo como objetivo simplificar los cálculos y eliminar anomalías, así como también, atributos recursivos, reducir el rango de valores y hacer que el proceso de formación sea más rápido.

Los sensores objetivos son un eje principal en este modelamiento debido a que estos sensores captan los datos referentes a las concentraciones de metano, los cuales son: MM263, MM256, MM264. Inicialmente se utilizaron los modelos SNN, LSTM, GRU y posteriormente con redes neuronales bidireccionales en donde se observó una mejora en la eficiencia del sistema de alerta respecto a las RN simples. Se utilizaron métricas para evaluar la sensibilidad, efectividad y puntaje con el fin de categorizar la fidelidad y éxito del modelo que más se relacione a los datos reales.

3.1 Resultados PCA

El PCA se aplicó con la finalidad de que la variación de datos sea explicada como mínimo en el 89% de la data. El PCA, en sí, usa la reducción de dimensiones con lo cual se pudo eliminar la correlación de varias variables y sustituir los indicadores originales con un número menor de componentes que pueden retener la información más representativa de estos datos. El resultado indicó una positiva correlación, es decir los valores de las variables utilizadas tienden a incrementarse de manera homóloga. La data es básicamente explicada en ocho componentes principales, el intervalo de confianza es marcado de color verde en la Figura 3.1 y denota los datos que están positivamente relacionados entre sí, los que están fuera del área son conocidos como datos erráticos. Los puntos observados en la Figura 3.1 forman parte de las correlaciones de conjuntos de las variables que se utilizaron para el modelamiento predictivo. Estos conjuntos están

agrupados por la correlación de datos existente, asegurando que al menos el 90% de estos datos tienen relación.



Figuras 3.1 Visualización de series temporales y PCA

3.2 Resultados RFE

Utilizando la eliminación recursiva de funciones, se inició el proyecto con 34 variables. Al trabajar con una gran cantidad de datos se observa que ciertas variables no tienen un peso importante en el modelo de predicción del modelo, por lo cual se seleccionó los atributos más relevantes y se restructuró de forma que mejor se acople al modelo. La data está conformada por filas y columnas (muestras y atributos). La selección de atributos consiste en elegir un subconjunto de variables que son relevantes entre el conjunto de datos. En la Tabla 3.1 se visualiza las columnas que se obtuvieron de la eliminación de atributos y funciones recursivas, estos atributos están relacionados e influyen directa o indirectamente en la concentración de metano y la calidad de aire.

Tabla 3.1 Relación de atributos y sensores con el Metano.

Columna	Sensor	Campo
6	AN311	Anemómetro
9	TP1721	Temperatura
11	BA1723	Barómetro
12	TP1711	Temperatura
13	RH1712	Humedad
17, 18, 19, 22, 23	Sensores	Métricas de Metano

Para la eliminación de las variables no necesarias se utilizó el método de Eliminación de Atributos Recursivos (EAR). Este método buscó un subconjunto de atributos analizando todas las variables y eliminando los peores resultados al final de cada iteración. Se generaron 110 procesos iterativos en donde se eliminaron las variables menos relacionadas de cada ciclo, generando un modelo de regresión. La ventaja de este método, además de escoger los mejores resultados, brinda un sin número de ciclos a utilizar según el criterio del programador. El resultado final es un conjunto de datos con menos variables que al inicio, relacionando las más importantes con la concentración de metano.

3.3 Función de auto correlación para múltiples sensores de metano

La función de auto correlación se aplicó a los sensores objetivos con el fin de encontrar y elaborar la ventana móvil que entrará en relación con todos los datos, donde se denotara un “Max Lag” que en sí es el tiempo límite en donde la correlación es máxima manteniéndose en dicho tiempo y en relación con los sensores objetivos. Se eligió gráficamente el menor tiempo de tendencia en correlación a los sensores, ya que este tiempo estaría contenido en los 3 sensores, la Figura 3.2 denotó la correlación máxima de los sensores, que en nuestro caso fue el sensor MM256.

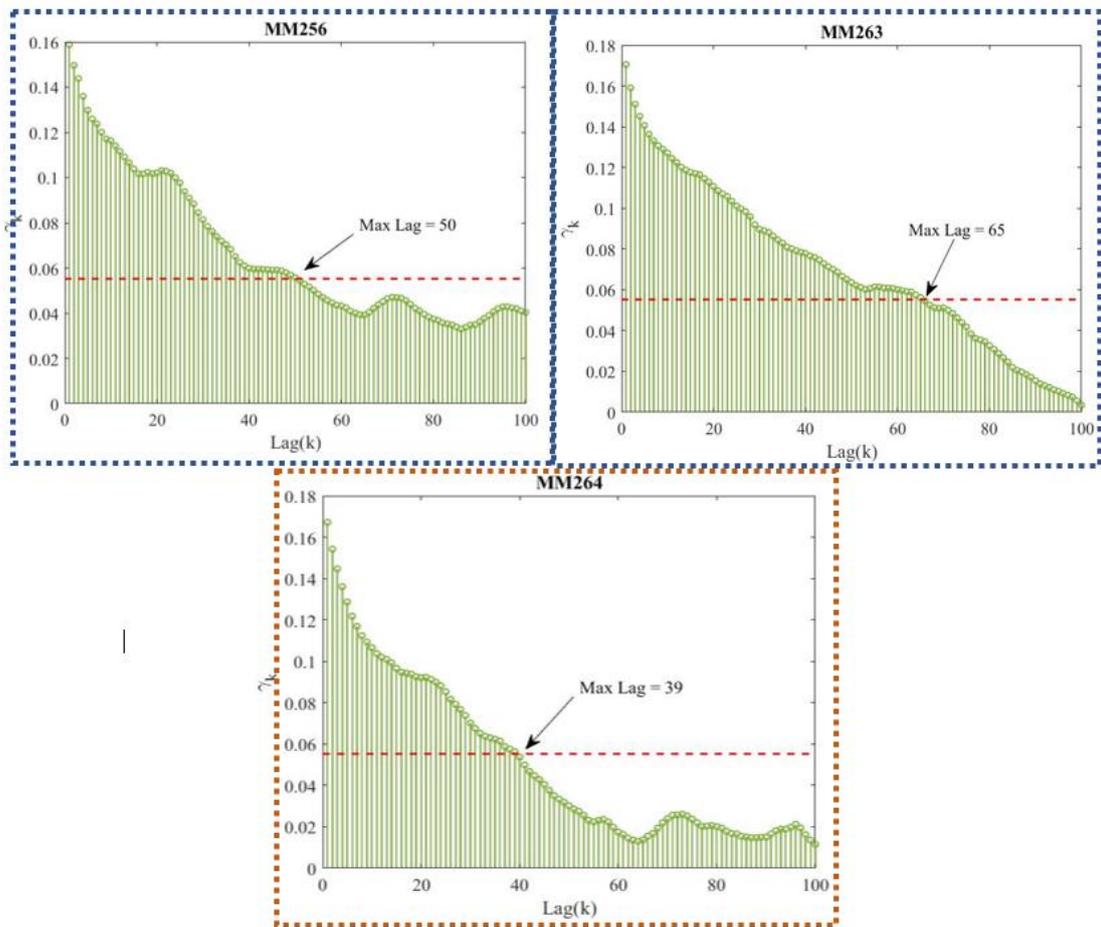


Figura 3.2 Función de auto correlación entre los sensores MM263, MM264, MM256.

El “Max Lag” se interpreta como un límite de datos altamente correlacionados, aquellos datos sobre la línea roja son los datos que más correlacionados están y los que se encuentra por debajo de la misma son los que menos relación tienen entre sí. Se utilizó una ventana móvil de 39 ciclos en donde los 3 sensores se relacionan entre sí y es necesario para utilizar toda la data de entrada separándola en estos ciclos, lo cual nos permite predecir la concentración de metano en tiempo futuro. Por ende, es necesario saber qué tan correlacionados están los datos y en qué medida. Este proceso solo se realiza para redes recurrentes con un número limitado de ciclos y variables que sirvieron para dividir la data en conjuntos pequeños para el entrenamiento del modelo.

3.4 Entrenamiento de modelos y optimización

Concretamos los procesos previos para la obtención de la mejora de los modelos como la limpieza de datos y eliminación de anomalías para mejorar los modelos de predicción. Se observó que las redes neuronales bidireccionales son modelos de predicción más precisos, según lo que se evaluó en las métricas que se observan que los datos predichos por el modelo Bidireccional LSTM se acerca más a los datos reales. En la Figura 3.3 se presenta el modelo keras optimizado utilizado en el estudio.

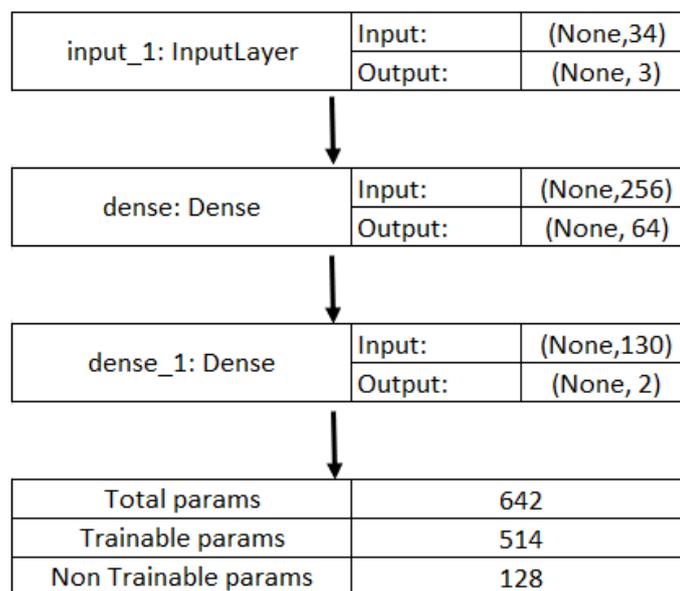


Figura 3.3 Modelamiento optimizado, función Keras.

a) Series Temporales: Modelos de Predicción

Se analizó cada modelo de predicción mediante un análisis gráfico, representando los datos predichos, así como también los datos reales en series temporales, además se realizó un análisis numérico evaluado según el índice de fiabilidad de las métricas utilizadas en el proyecto. Las Figuras 3.4 y 3.5 denotan los modelos evaluados. Las Figuras representan gráficamente la predicción (rojo) en relación con los valores reales (azul), ambos modelos están basados en redes neuronales recurrentes

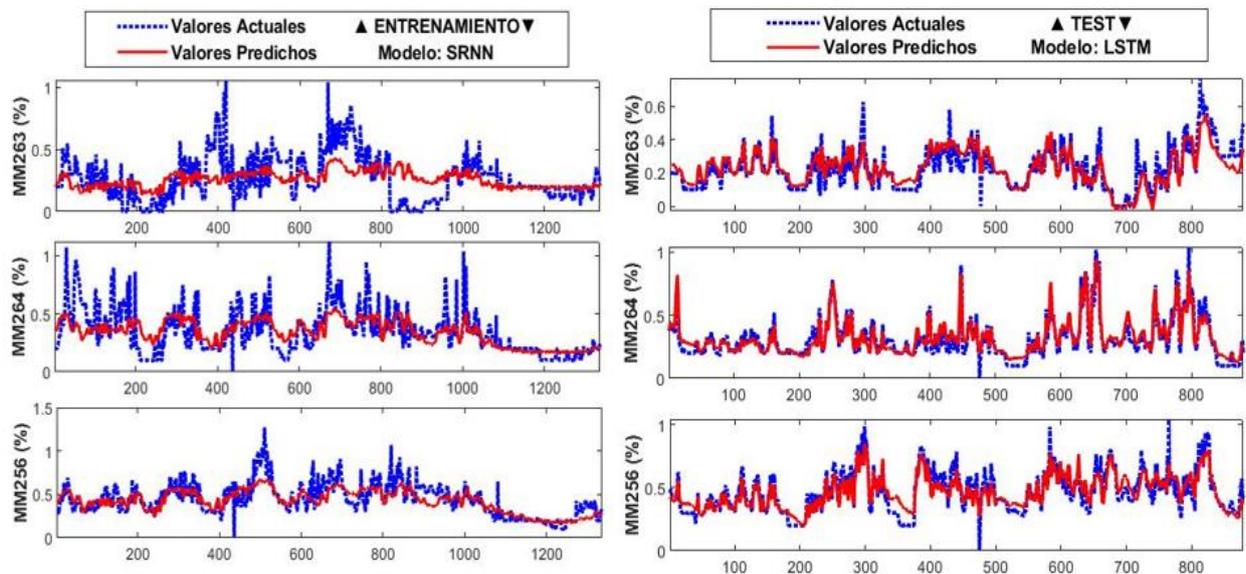


Figura 3.4 Series temporales: Modelos SRNN, LSTM.

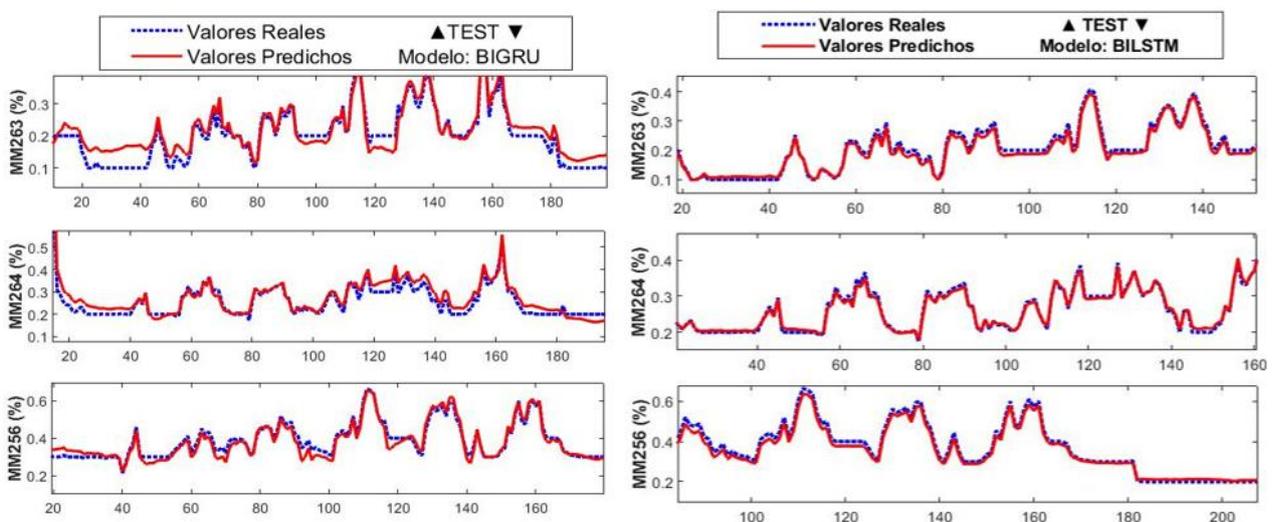


Figura 3.5 Series temporales: Modelos BIGRU, BiLSTM.

La Figura 3.4 así mismo, representa gráficamente la predicción (rojo) en relación con los valores reales (azul), ambos modelos están basados en redes neuronales recurrentes bidireccionales. Es notorio que entre los modelos de predicción el que mejor correlacionado se encuentra con la realidad es el modelo BiLSTM, sin embargo, el puntaje obtenido por las métricas que se utilizaron en el proyecto nos brindara una mejor idea del porcentaje de fiabilidad de dicho modelo.

b) Resultados Red Neuronal BILSTM

Los resultados visibles del modelo neuronal BILSTM son una serie de configuraciones que utilizo una propagación interna a corto plazo, en la Figura 3.5 se puede constatar como la observación gráfica y numérica de las métricas coincide con los diagramas de dispersión es decir en todos los casos los valores predichos son muy cercanos a la realidad.

El diagrama de dispersión que se observa en la anterior figura relaciono los datos reales con los predichos de una manera lineal, denotando la gran relación que estos tienen. En la siguiente Figura 3.6 se observan las correlaciones lineales por diagramas de dispersión de los otros modelos de predicción, demostrando así que el modelo BILSTM predice y brinda los mejores resultados en comparación a todos los modelos.

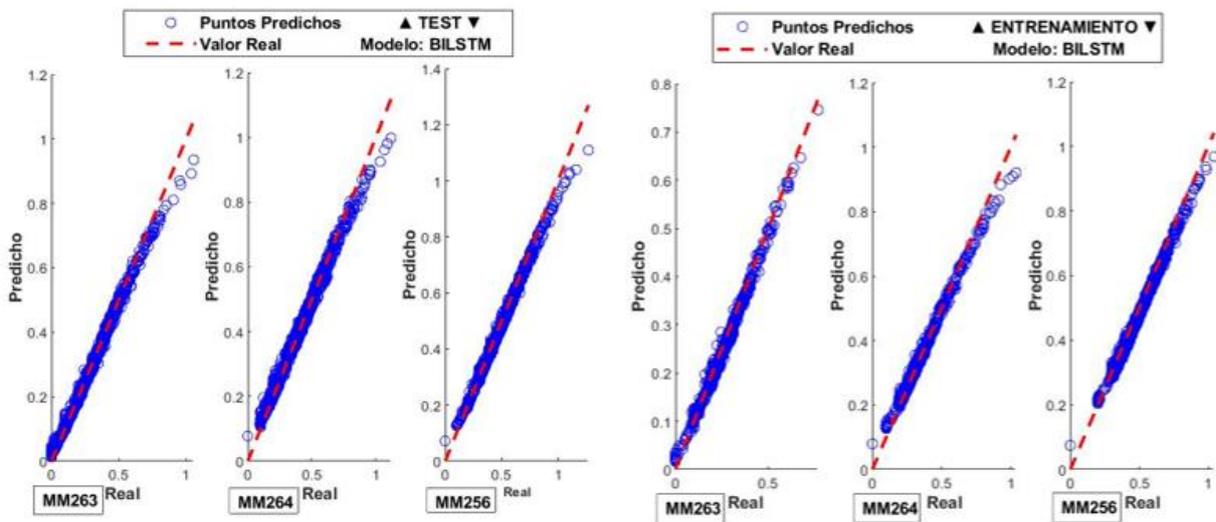


Figura 3.6 Diagrama de dispersión de la predicción de datos del entrenamiento/test en el modelo BILSTM.

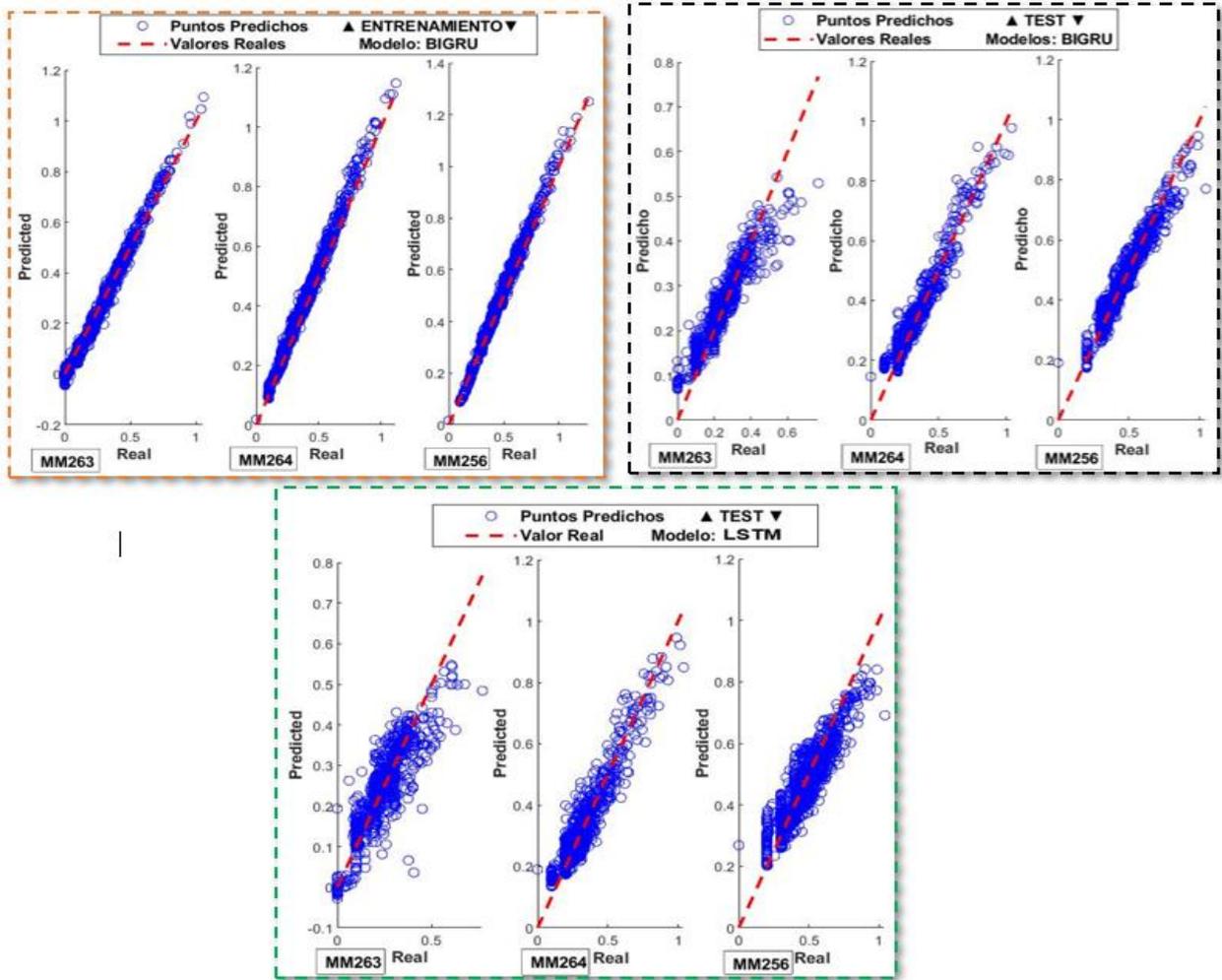


Figura 3.7 Diagrama de dispersión modelos LSTM, BiGRU

Con los valores residuales absolutos del modelo, se comparó los resultados presentados en la Figura 3.7, el mejor resultado de los 3 sensores estudiados lo obtuvo el MM264. La información nos indica cuál de estos sensores se ajusta mejor a la data, tomando en cuenta que el mejor ajuste del modelo se obtiene cuando la serie temporal está cercana al origen.

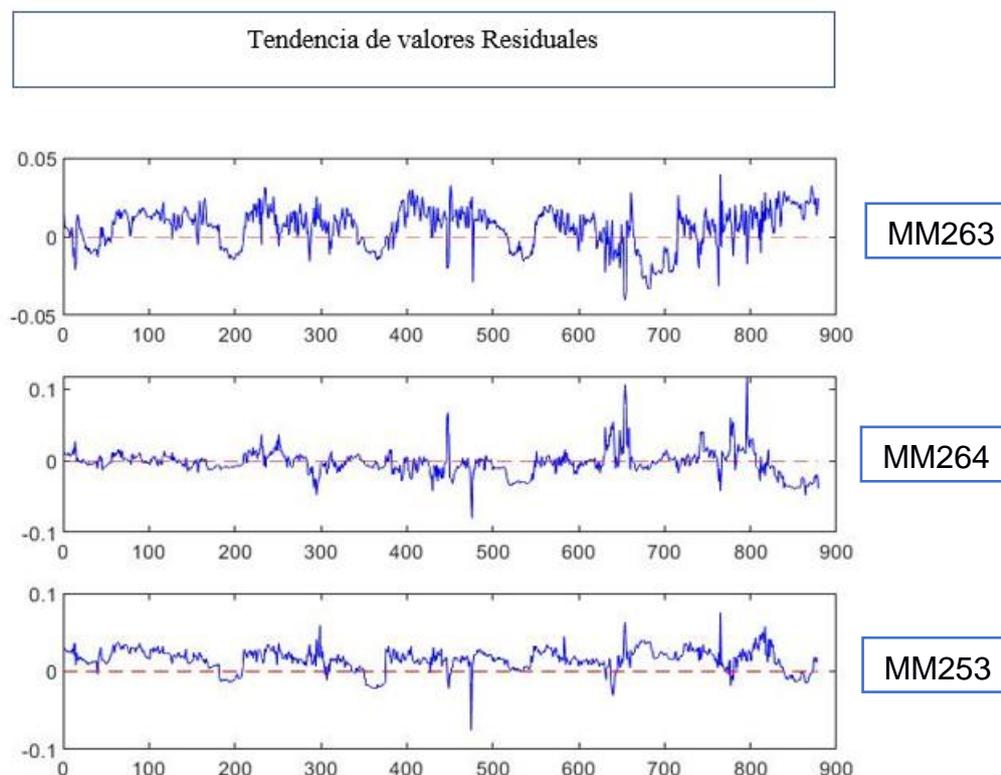


Figura 3.8 Valores residuales de sensores objetivos en el modelo BILSTM

3.5 Evaluación de modelos predictivos

Los diagramas de dispersión, así como las gráficas de series temporales anteriormente visualizadas fueron utilizadas con el fin de reconocer y corroborar gráficamente el modelo que mejor se acople a la información real obtenida del conjunto de datos, siendo el mejor modelo el BILSTM (red neuronal recurrente bidireccional a corto plazo). Un reconocimiento numérico fue necesario para complementar el resultado final, haciendo referencia a las métricas RMSE y RR presentes en la Tabla 2.2. Adicional a este análisis numérico, los valores residuales de los modelos fueron utilizados para elaborar un diagrama BoxPlot con el fin de recabar más información gráfica sobre el mejor modelo y comparar los que mejores resultados brindaron. En la siguiente Figura 3.9, se denota que los modelos con mejor predicción para concentraciones de gas metano en minería subterránea de carbón son los modelos bidireccionales seguido del modelo GRU, LSTM y SRNN.

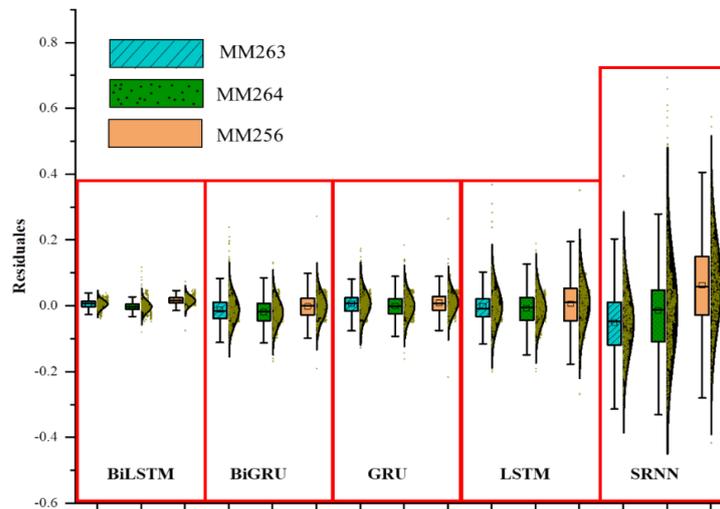


Figura 3.9 Comparación de Modelos Predictivos Recurrentes.

Se observa que el modelo con mejores resultados es el modelo BiLSTM debido a que los datos residuales obtenidos de este modelo presentan una casi perfecta similitud con los datos reales.

Tabla 3.2 Evaluación del puntaje por Modelo utilizado.

Modelos Predictivos		Métrica	
		RMSE	RR
SRRN	Entrenamiento	0.4021+/-0.2296	0.1520+/-0.0456
	Test	0.2072+/- 0.0686	0.2072+/-0.0687
LSTM	Entrenamiento	0.0460+/-0.0035	0.9537+/-0.0057
	Test	0.0598+/-0.0092	0.8273+/-0.0646
GRU	Entrenamiento	0.0242+/-0.0043	0.9872+/-0.0029
BiLSTM	Entrenamiento	0.0182+/-0.0026	0.9949+/-0.0014
	Test	0.0172+/-0.0035	0.9938+/-0.0014
BiGRU	Entrenamiento	0.0256+/-0.0051	0.9921+/-0.0007
	Test	0.0434+/-0.0026	0.9228+/-0.0253

El modelo BILSTM generó un 99.50 % de efectividad en su predicción, mientras que el modelo BIGRU un 99.20%; ambos modelos denotan un alto índice de efectividad para la predicción, sin embargo, existe una gran diferencia gráfica entre estos. Es importante generar el análisis gráfico y numérico de los modelos con el fin de observar el que mejor predicción realizó entre estos modelos. El modelo GRU se puntuó también un 98% seguido del modelo LSTM con 88.5% en promedio de ambas métricas, el peor resultado lo obtuvo el modelo SRNN con un 20% de fiabilidad en las métricas, lo cual se puede corroborar gráficamente en la serie temporal del modelo.

CAPÍTULO 4

4. CONCLUSIONES Y RECOMENDACIONES

4.1 Conclusiones y trabajo futuro

En este proyecto se planteó diseñar un modelo de predicción de la concentración de metano presente en labores mineras subterráneas de carbón. Para lo cual, se evaluó y utilizó un conjunto inicial de 9 millones de datos. Con las redes recurrentes y bidireccionales utilizadas, se planteó resolver problemas desafiantes en la industria del carbón, a través de la medición y predicción del gas metano presente en las labores de mina subterránea.

Se propone el uso de funciones de pérdida de información irrelevante, con la finalidad de aumentar la eficacia del modelo al reducir el peso de la información procesada. En esta etapa, el entrenamiento y la evaluación de redes es indispensable para el éxito del proyecto; se corrieron los modelos LSTM, GRU, RRNN, basados en redes neuronales recurrentes, también se utilizó redes neuronales bidireccionales (BILSTM Y BIGRU). Los modelos bidireccionales obtuvieron los mejores resultados con una confiabilidad de predicción del 99.49% y 98.20, respectivamente.

La revisión bibliográfica de las diferentes técnicas planteadas en este proyecto fue utilizada para determinar las limitaciones y ventajas de este, así como también, destaca la capacidad de predicción de una red neuronal y el potencial para aplicar dichos modelos a otros problemas en la industria minera. En este proyecto fue necesario implementar varios modelos que pronostiquen la concentración de metano y compararlos, para seleccionar el que mejor se acople a la información dando los mayores resultados.

Abordar el problema de predecir el porcentaje de metano presente en varias zonas de la mina fue la meta principal para la inserción de IA en minería; se propuso un esquema supervisado donde las predicciones de concentraciones de metano captadas por los sensores objetivos son evaluadas con el fin de comparar los mejores resultados. En este contexto, los modelos como SRNN, LSTM Y GRU no obtuvieron gran éxito. Sin embargo, para superar la limitación de estos modelos se propuso un enfoque basado en redes neuronales bidireccionales corroborando los resultados gráficos con los numéricos a través de métricas.

Se diseñó un algoritmo que combinó una ventana móvil multidireccional para el proceso de predicción y una arquitectura de red neuronal con múltiples entradas, capas convolucionales de una dimensión, capas Long Short Time Memory LSTM y capas Dense. Los modelos BIGRU y GRU tuvieron los mejores resultados luego del modelo BISLTM. Estos modelos contemplan la utilización de inteligencia artificial para el desarrollo del monitoreo de operaciones mineras.

El modelo BILSTM demuestra que la programación inteligente puede extraer características relacionadas con la predicción de metano de manera más efectiva que los humanos. El aprendizaje profundo demuestra ser de un gran potencial para ser aplicado en minería subterránea del carbón y no solo en esta rama, en sí, podremos aplicar el modelo a distintos problemas en procesos mineros.

Como trabajo futuro, el modelamiento de series temporales basadas en diferentes arquitecturas de aprendizaje profundo es necesario para mejorar la efectividad de la predicción debido a que le permite al modelo actualizar datos y reducir pesos para un mejor procesamiento de la información, siendo este un enfoque ideal del proyecto. Así como también, se deben considerar nuevas técnicas de normalización y un mejor preprocesamiento de datos para optimizar los resultados ya obtenidos. Los codificadores automáticos han demostrado resultados interesantes para la mejora de información obtenida por sensores.

El incremento del número de datos para el entrenamiento de la red puede mejorar la capacidad de predicción. La estimación de concentraciones de gas metano es crucial para la minería del carbón, debido a su potencial de predicción y prevención de accidentes, posibilitando la identificación temprana de accidentes que pueda afectar al desarrollo minero.

4.2 Recomendaciones

El evaluar varios modelos ayuda a maximizar el porcentaje de éxito en la predicción de la concentración de metano, esto debido a que entre más eficaz el modelo, mayor sería

la sensibilidad para alertar a los operarios de una mina sobre la concentración alarmante en su entorno.

Se recomienda seguir entrenando el modelo con nueva información proveniente de un nuevo flujo de datos en distintos meses del año, así como también en casos de desprendimientos instantáneos de gas metano que pueden ocurrir mientras las operaciones mineras influyen en las presiones de los estratos. En esta línea, se recomienda repetir el tamaño de muestra para confirmar las conclusiones, así como también para probar nuevas combinaciones o incluir nueva metodología de pronóstico.

BIBLIOGRAFÍA

- Ait-Amir, B., Pougnet, P., & El Hami, A. (2015). Meta-Model Development. *Embedded Mechatronic Systems*, 2, 151–179. <https://doi.org/10.1016/B978-1-78548-014-0.50006-2>
- Asheri Arnon, T., Ezra, S., & Fishbain, B. (2019). Water characterization and early contamination detection in highly varying stochastic background water, based on Machine Learning methodology for processing real-time UV-Spectrophotometry. *Water Research*, 155, 333–342. <https://doi.org/10.1016/J.WATRES.2019.02.027>
- Bonet Cruz, I., Salazar Martinez, S., Abed, A. R., Abalo, G., & Lorenzo, M. M. G. (n.d.). *Redes neuronales recurrentes para el análisis de secuencias Recurrent neural network for sequences analysis*.
- Ciesielczuk, J., Górka, M., Fabiańska, M. J., Misz-Kennan, M., & Jura, D. (2021). The influence of heating on the carbon isotope composition, organic geochemistry and petrology of coal from the Upper Silesian Coal Basin (Poland): An experimental and field study. *International Journal of Coal Geology*, 241. <https://doi.org/10.1016/J.COAL.2021.103749>
- Damian, D. (2015). *Spyder, un potente entorno de desarrollo interactivo para Python | Ubunlog*. <https://ubunlog.com/spyder-entorno-desarrollo-python/>
- Dubinski, Józ., Stec, K., & Bukowska, M. (2019). Geomechanical and tectonophysical conditions of mining-induced seismicity in the upper silesian coal basin in Poland: A case study. *Archives of Mining Sciences*, 64(1), 163–180. <https://doi.org/10.24425/ams.2019.126278>
- Eshaghian, O., Hoseinie, S. H., & Maleki, A. (2021). Multi-attribute failure analysis of coal cutting picks on longwall shearer machine. *Engineering Failure Analysis*, 120, 105069. <https://doi.org/10.1016/J.ENGFANAL.2020.105069>
- Jason, F. (2018). *Eliminación recursiva de funciones (RFE) para la selección de funciones en Python*. 2018. <https://machinelearningmastery.com/rfe-feature-selection-in-python/>
- Kang, B., & Choo, H. (2016). A deep-learning-based emergency alert system. *ICT Express*, 2(2), 67–70. <https://doi.org/10.1016/J.ICTE.2016.05.001>
- Kotarba, M. J., Słowakiewicz, M., Misz-Kennan, M., Więclaw, D., Jurek, K., & Waliczek, M. (2021). Simulated maturation by hydrous pyrolysis of bituminous coals and carbonaceous shales from the Upper Silesian and Lublin basins (Poland): Induced compositional variations in biomarkers, carbon isotopes and macerals. *International Journal of Coal Geology*, 247. <https://doi.org/10.1016/J.COAL.2021.103856>
- Kozielski, M., Sikora, M., & Wróbel, Ł. (2021). Data on methane concentration collected by underground coal mine sensors. *Data in Brief*, 39, 107457. <https://doi.org/10.1016/J.DIB.2021.107457>
- Kumaraswamy, B. (2021). Neural networks for data classification. *Artificial Intelligence in Data Mining: Theories and Applications*, 109–131. <https://doi.org/10.1016/B978-0-12-820601-0.00011-2>

- Lasso, A., & Kazanzides, P. (2019). System integration. *Handbook of Medical Image Computing and Computer Assisted Intervention*, 861–891. <https://doi.org/10.1016/B978-0-12-816176-0.00040-5>
- Li, J., Lu, G., Niu, T., & Zhang, J. (2020). Developing an online air quality warning system based on streaming data for dynamic environmental management. *Journal of Cleaner Production*, 273, 122953. <https://doi.org/10.1016/j.jclepro.2020.122953>
- Loy-Benitez, J., Heo, S. K., & Yoo, C. K. (2020). Soft sensor validation for monitoring and resilient control of sequential subway indoor air quality through memory-gated recurrent neural networks-based autoencoders. *Control Engineering Practice*, 97, 104330. <https://doi.org/10.1016/J.CONENGPRAC.2020.104330>
- Loy-Benitez, J., Tariq, S., Nguyen, H. T., Safder, U., Nam, K. J., & Yoo, C. K. (2022). Neural circuit policies-based temporal flexible soft-sensor modeling of subway PM2.5 with applications on indoor air quality management. *Building and Environment*, 207, 108537. <https://doi.org/10.1016/J.BUILDENV.2021.108537>
- Loy-Benitez, J., Vilela, P., Li, Q., & Yoo, C. K. (2019). Sequential prediction of quantitative health risk assessment for the fine particulate matter in an underground facility using deep recurrent neural networks. *Ecotoxicology and Environmental Safety*, 169(August 2018), 316–324. <https://doi.org/10.1016/j.ecoenv.2018.11.024>
- Mariño, J. E., Castro, A., Botero, A. Á., Mojica, L., Granados, D. P., & Acuña, C. M. (2015). *Gas asociado al carbón (CBM o GMAC). Geología, contenidos, reservas, minería y posibilidades en Colombia.*
- Mendecki, M. J., Szczygieł, J., Lizurek, G., & Teper, L. (2020). Mining-triggered seismicity governed by a fold hinge zone: The Upper Silesian Coal Basin, Poland. *Engineering Geology*, 274, 105728. <https://doi.org/10.1016/J.ENGGEOL.2020.105728>
- Na, J., Jeon, K., & Lee, W. B. (2018). Toxic gas release modeling for real-time analysis using variational autoencoder with convolutional neural networks. *Chemical Engineering Science*, 181, 68–78. <https://doi.org/10.1016/J.CES.2018.02.008>
- Neira, I. N. G. L. (2015). *UNIVERSIDAD POLITÉCNICA SALESIANA SEDE GUAYAQUIL CARRERA : INGENIERÍA ELECTRÓNICA Proyecto técnico previo a la obtención del título de : INGENIERO ELECTRÓNICO Título : “ DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA DE MAQUETA DE PLANTA DE TRATAMIENTO DE AGUAS .”*
- Nguyen, T. P., Yeh, C. T., Cho, M. Y., Chang, C. L., & Chen, M. J. (2022). Convolutional neural network bidirectional long short-term memory to online classify the distribution insulator leakage currents. *Electric Power Systems Research*, 208, 107923. <https://doi.org/10.1016/J.EPSR.2022.107923>
- Ocaña-Fernández, Y., Alex Valenzuela-Fernández, L., & Lourdes Garro-Aburto, L. (n.d.). *Inteligencia artificial y sus implicaciones en la educación superior Artificial Intelligence and its Implications in Higher Education*. 7, 536–568. <https://doi.org/10.20511/pyr2019.v7n2.274>
- Sharma, D. K., Chatterjee, M., Kaur, G., & Vavilala, S. (2022). Deep learning applications for disease diagnosis. *Deep Learning for Medical Applications with Unique Data*, 31–51. <https://doi.org/10.1016/B978-0-12-824145-5.00005-8>
- Wang, J., Ji, J., Ravikumar, A. P., Savarese, S., & Brandt, A. R. (2022). VideoGasNet:

Deep learning for natural gas methane leak classification using an infrared camera. *Energy*, 238, 121516. <https://doi.org/10.1016/J.ENERGY.2021.121516>

Yuan, L., Jiang, Y., He, X., Dou, L., Zhao, Y., Zhao, X., Wang, K., Yu, Q., Lu, X., & Li, H. (2018). Research progress of precise risk accurate identification and monitoring early warning on typical dynamic disasters in coal mine. In *Meitan Xuebao/Journal of the China Coal Society* (Vol. 43, Issue 2, pp. 306–318). China Coal Society. <https://doi.org/10.13225/j.cnki.jccs.2017.4151>

ANEXOS

Anexo A

Modelo Predictivo SRNN					
Training			Test		
	RMSE	RR		RMSE	RR
MM263	0.1623	0.2019	MM263	0.2489	0.1148
MM264	0.4246	0.1417	MM264	0.1281	0.1527
MM256	0.6196	0.1124	MM256	0.2447	0.1528
Promedio	0.4021	0.1522	Promedio	0.2072	0.1401
Error %	0.2296	0.0454	Error %	0.0686	0.0219
Modelo Predictivo LSTM					
Training			Test		
	RMSE	RR		RMSE	RR
MM263	0.0292	0.9839	MM263	0.0392	0.9152
MM264	0.0226	0.9886	MM264	0.0381	0.9441
MM256	0.021	0.9893	MM256	0.0356	0.9536
Promedio	0.0242	0.9872	Promedio	0.0376	0.9376
Error %	0.0043	0.0029	Error %	0.0018	0.0199
Modelo Predictivo GRU					
Training			Test		
	RMSE	RR		RMSE	RR
MM263	0.0421	0.9486	MM263	0.0577	0.7642
MM264	0.0472	0.9534	MM264	0.0518	0.8933
MM256	0.0488	0.9601	MM256	0.0699	0.8244
Promedio	0.0462	0.9536	Promedio	0.0598	0.8273
Error %	0.0035	0.0057	Error %	0.0092	0.0645
Modelo Predictivo BILSTM					
Training			Test		
	RMSE	RR		RMSE	RR
MM263	0.0161	0.9955	MM263	0.0137	0.9922
MM264	0.0211	0.9933	MM264	0.0174	0.9947
MM256	0.0176	0.9961	MM256	0.0207	0.9946
Promedio	0.0182	0.9949	Promedio	0.0172	0.9938
Error %	0.0026	0.0014	Error %	0.0035	0.0014
Modelo Predictivo BIGRU					
Training			Test		
	RMSE	RR		RMSE	RR
MM263	0.0204	0.9915	MM263	0.0456	0.8936
MM264	0.0307	0.9921	MM264	0.0442	0.9365
MM256	0.0257	0.9929	MM256	0.0405	0.9383
Promedio	0.0256	0.9921	Promedio	0.0434	0.9228
Error %	0.0051	0.0007	Error %	0.0026	0.0253

En el anexo adjunto se puede observar los resultados numéricos de los modelos evaluados, denotando el mejor resultado del modelo BILSTM, con un resultado en la métrica RR de 99.48%.

Anexo B

Selección de atributos vía RFE

```
from sklearn.feature_selection import RFE #Importar clase desde sklearn
from sklearn.tree import DecisionTreeRegressor
from sklearn.pipeline import Pipeline
from numpy import mean, std

#Primero designamos un estimador (OTRO modelo ML) y asignamos el número de
atributos a seleccionar.

rfe = RFE(estimator = DecisionTreeRegressor(), n_features_to_select = 11)

model = DecisionTreeRegressor() #realiza un modelo igual al estimador

#Crear el pipeline
pipeline = Pipeline(steps = [('s',rfe),('m', model)])

#Dividir la data en diferentes splits y asignar repetición de experimento
cv = RepeatedKFold(n_splits = 10, n_repeats = 11)

#Cuantificar los scores de importancia
n_scores = cross_val_score(pipeline, X_train, y_train,
                           scoring='neg_mean_absolute_error',
                           cv=cv,n_jobs = -1,
                           error_score='raise')

# Reporte de performance
print('MAE:      %.3f      (%.3f)'      %      (mean(n_scores),      std(n_scores)))

rfe.fit(X, Y)

print("##### ~Recursive Feature Elimination Ranking~ #####")
```

Se presenta la programación de la función de eliminación de atributos recursivos, utilizada en el preprocesamiento de datos.

Anexo C

Uso de métricas de evaluación RMSE, RR.

```
def rmse(predictions, targets):
    return np.sqrt(((predictions - targets) ** 2).mean ())

import numpy as np
rmse_lr = np.zeros(2) # [0 0]

for i in range(2):
    rmse_lr[i]= rmse(yhat[:,i],y_test[:,i]) # [0.3 0.5]

av_rmse_lr = np.mean(rmse_lr) #Promedio
un_lr_rmse = av_rmse_lr/np.sqrt(2) #Generamos incertidumbre

plt.figure(dpi =250)
plt.subplot(211)
plt.plot(yhat[:,0], 'r', label = 'Predicted')
plt.plot(y_test[:,0], 'b--', label = 'Actual')
plt.legend()
plt.subplot(212)
plt.plot(yhat[:,1], 'r', label = 'Predicted')
plt.plot(y_test[:,1], 'b--', label = 'Actual')
#plt.legend(loc='best')

plt.figure(dpi = 250)
plt.subplot(211)
plt.scatter(y_test[:,0],yhat[:,0],c='b',marker = '*')
plt.plot([min(y_test[:,0]), max(y_test[:,0])],[min(yhat[:,0]), max(yhat[:,0])], color = 'r',
linewidth =2, linestyle = '--', label = 'Reference line')
plt.xlabel("Actual")
plt.ylabel("Predicted")
plt.legend()

plt.subplot(212)
plt.scatter(y_test[:,1],yhat[:,1],c='b',marker = '*')
plt.plot([min(y_test[:,1]), max(y_test[:,1])],[min(yhat[:,1]), max(yhat[:,1])], color = 'r',
linewidth =2, linestyle = '--', label = 'Reference line')
plt.xlabel("Actual")
plt.ylabel("Predicted")
```