

**UNIVERSIDAD DE CONCEPCIÓN - CHILE  
FACULTAD DE INGENIERÍA  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

*“Proyecciones de producción de cobre considerando consumo de agua, energía y emisiones de gases de efecto invernadero, con redes neuronales recurrentes”*

por

**Pablo Sebastián Contreras Cerpa**



Profesora Guía  
**Dra. Lorena Pradenas Rojas**

Concepción, Agosto de 2021

Tesis presentada a la

**DIRECCIÓN DE POSTGRADO  
UNIVERSIDAD DE CONCEPCIÓN**



Para optar al grado de

**MAGISTER EN INGENIERÍA INDUSTRIAL**

© 2021 Pablo Sebastián Contreras Cerpa

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.



# ÍNDICE GENERAL

ÍNDICE GENERAL .....	ii
ÍNDICE DE TABLAS .....	iv
ÍNDICE DE FIGURAS .....	v
1. INTRODUCCIÓN.....	1
1.1. Hipótesis.....	2
1.2. Objetivos .....	2
1.2.1. Objetivo general.....	2
1.2.2. Objetivos específicos.....	2
1.3. Estructura del informe .....	2
2. CONTEXTO.....	3
2.1. Producción de cobre en el mundo .....	3
2.2. Producción de cobre en Chile .....	4
2.3. Insumos críticos para la producción de cobre en Chile .....	5
2.3.1. Consumo de agua.....	5
2.3.2. Consumo de electricidad.....	6
2.3.3. Emisiones de gases de efecto invernadero .....	7
3. MARCO TEÓRICO.....	8
3.1. Machine Learning.....	8
3.2. Redes neuronales artificiales .....	8
3.2.1. Redes neuronales recurrentes .....	10
3.2.2. Red LSTM.....	12
3.2.3. Red GRU .....	13
3.3. Series de tiempo .....	15
4. REVISIÓN DE LA LITERATURA .....	16
4.1. Redes neuronales y las proyecciones de producción .....	16
4.2. Redes neuronales recurrentes y las proyecciones de producción.....	17
4.3. Brecha del conocimiento .....	19

5. METODOLOGÍA DE ESTUDIO .....	20
5.1. Conjunto de datos.....	20
5.1.1. Producción de Cobre.....	21
5.1.2. Matriz de correlación datos de producción de cobre .....	22
5.2. Modelos .....	23
6. RESULTADOS .....	26
6.1. Resultados modelo RNR.....	26
6.2. Resultados modelo GRU .....	28
6.3. Resultados modelo LSTM.....	29
6.4. Discusión .....	31
7. CONCLUSIONES .....	33
REFERENCIAS.....	35
A. ANEXOS .....	38
A.1 Anexos modelo RNR .....	38
A.2 Anexos modelo GRU .....	40
A.3 Anexos modelo LSTM.....	42



## ÍNDICE DE TABLAS

Tabla 2.1: Producción mundial de cobre en miles de toneladas entre los años 2018-2019 [5].	3
Tabla 2.2. Indicadores económicos de producción y exportación de cobre entre los años 2010-2019 [6].	4
Tabla 4.1. Resumen de estudios disponibles en la literatura especializada	19
Tabla 5.1. Matriz de datos para producción de cobre.	20
Tabla 5.2. Parámetros de configuración experimentales de los modelos RNR, GRU y LSTM.	25
Tabla 6.1. Resultados de métricas: MAE, MAPE, MSE y $R^2$ en predicción de producción de cobre con modelo RNR.	27
Tabla 6.2. Resultados de métricas: MAE, MAPE, MSE y $R^2$ en predicción de producción de cobre con modelo GRU.	28
Tabla 6.3. Resultados de métricas: MAE, MAPE, MSE y $R^2$ en predicción de producción de cobre con modelo LSTM.	29



## ÍNDICE DE FIGURAS

Figura 3.1. Estructura de una Red Neuronal Artificial [elaboración propia].....	9
Figura 3.2. Estructura Red Neuronal Recurrente [elaboración propia]. .....	10
Figura 3.3. Estructura Red Neuronal Recurrente tipo LSTM [elaboración propia]. .....	12
Figura 3.4. Estructura Red Neuronal Recurrente tipo GRU [elaboración propia].....	14
Figura 5.1. Producción de cobre mensual en Chile, años 2000-2020 [elaboración propia]. .....	21
Figura 5.2. Histograma producción de cobre en Chile [elaboración propia]. .....	21
Figura 5.3. Matriz de correlación datos producción de cobre [elaboración propia]. .....	22
Figura 5.4. Metodología para los modelos RNR, GRU y LSTM [elaboración propia]. .....	23
Figura 6.1. Gráfico de predicción de producción de cobre del modelo RNR (64-64) vs producción real. .....	27
Figura 6.2. Gráfico de predicción de producción de cobre del modelo GRU (64-64) vs producción real. ....	29
Figura 6.3. Gráfico de predicción de producción de cobre del modelo LSTM (64-64) vs producción real. ....	30
Figura 6.4. Grafica predicción modelos RNR, GRU y LSTM vs producción real.....	30
Figura A.1 Grafica predicción modelo RNR (96) vs producción real. ....	38
Figura A.2. Grafica predicción modelo RNR (64-96) vs producción real.....	38
Figura A.3. Grafica predicción modelo RNR (64-96-96) vs producción real.....	38
Figura A.4. Grafica entrenamiento modelo RNR (96) vs datos de entrenamiento.....	39
Figura A.5. Grafica entrenamiento modelo RNR (64-96) vs datos de entrenamiento. ....	39
Figura A.6. Grafica entrenamiento modelo RNR (64-96-96) vs datos de entrenamiento. ....	39
Figura A.7. Grafica predicción modelo GRU (96) vs producción real.....	40
Figura A.8. Grafica predicción modelo GRU (64-64) vs producción real. ....	40
Figura A.9. Grafica predicción modelo GRU (64-96-96) vs producción real. ....	40
Figura A.10. Grafica entrenamiento modelo GRU (96) vs datos de entrenamiento. ....	41
Figura A.11. Grafica entrenamiento modelo GRU (64-64) vs datos de entrenamiento. ....	41
Figura A.12. Grafica entrenamiento modelo GRU (64-96-96) vs datos de entrenamiento.....	41
Figura A.13. Grafica predicción modelo LSTM (96) vs producción real. ....	42

Figura A.14. Grafica predicción modelo LSTM (64-64) vs producción real..... 42

Figura A.15. Grafica predicción modelo LSTM (64-96-96) vs producción real..... 42

Figura A.16. Grafica entrenamiento modelo LSTM (96) vs datos de entrenamiento..... 43

Figura A.17. Grafica entrenamiento modelo LSTM (64-64) vs datos de entrenamiento. .... 43

Figura A.18. Grafica entrenamiento modelo LSTM (64-96-96) vs datos de entrenamiento. .... 43



## RESUMEN

La minería del cobre en Chile, es la base de la economía nacional y principal impulsor del desarrollo y la innovación del país. La importancia estratégica de la industria del cobre dirige a los principales agentes a la realización de estudios que les permitan mejorar permanentemente sus sistemas productivos, comerciales y de financiación, para mantener un adecuado equilibrio entre los aspectos estructurales de la actividad. A esto se le incluyen nuevos desafíos como: mayor dificultad para acceder al mineral, regulaciones ambientales más estrictas y la necesidad de optimizar los insumos críticos para su producción. Estos desafíos permitirán desarrollar una minería sustentable para el negocio y el medio ambiente. Debido a lo anterior, escasas investigaciones han estudiado estos nuevos desafíos para anticipar el comportamiento del negocio frente a distintos escenarios que puedan presentarse. Por otro lado, las herramientas computacionales han demostrado gran viabilidad para la realización de proyecciones permitiendo, mejorar la gestión de procesos y la planificación de recursos. En el presente estudio se proponen tres modelos basados en técnicas de redes neuronales artificiales del tipo recurrentes para proyectar la producción de cobre considerando: consumo de agua, consumo de energía y emisiones de gases de efecto invernadero. Estos modelos son: Red Neuronal Recurrente Simple, Red GRU y Red LSTM. Para entrenar y validar las redes se utilizaron datos reales de reportes entregados por el Ministerio de Minería y el Banco Central de Chile. Los resultados mostraron que el mejor resultado fue obtenido por el modelo de red LSTM, obteniendo las mejores métricas y un correcto ajuste a los datos. Luego, lo siguieron el modelo de Red GRU y la red neuronal recurrente simple. La importancia de este estudio radica en planificar planes de producción de cobre, considerando la sustentabilidad de la minería y el medio ambiente.

## **ABSTRACT**

Copper mining in Chile is the basis of the national economy and the main driver of the country's development and innovation. The strategic importance of the copper industry leads the main agents to carry out studies that allow them to permanently improve their production, commercial and financing systems, in order to maintain an adequate balance between the structural aspects of the activity. New challenges are added to this, such as: greater difficulty in accessing ore, stricter environmental regulations and the need to optimize critical inputs for its production. These challenges will make it possible to develop mining that is sustainable for the business and the environment. Due to the above, little research has studied these new challenges in order to anticipate the behavior of the business in the face of different scenarios that may arise. On the other hand, computational tools have demonstrated great viability for the realization of projections allowing to improve process management and resource planning. In the present study, three models based on artificial neural network techniques of the recurrent type are proposed to project copper production considering: water consumption, energy consumption and greenhouse gas emissions. These models are: Simple Recurrent Neural Network, GRU Network and LSTM Network. To train and validate the networks, real data from reports provided by the Ministry of Mining and the Central Bank of Chile were used. The results showed that the best result was obtained by the LSTM network model, obtaining the best metrics and a correct fit to the data. It was followed by the GRU network model and the simple recurrent neural network. The importance of this study lies in planning copper production plans, considering the sustainability of mining and the environment.

## GLOSARIO

**RNA:** Red Neuronal Artificial

**RNM:** Red Neuronal Multicapa

**RNR:** Red Neuronal Recurrente

**RNRM:** Red Neuronal Recurrente Multicapa

**LSTM:** *Long Short Term Memory*

**GRU:** *Gated Recurrent Unit*

**GEI:** Gases de Efecto Invernadero

**MSE:** *Mean Square Error*

**MAE:** *Mean Absolute Error*

**MAPE:** *Mean Absolute Percentage Error*

**R<sup>2</sup>:** Coeficiente de Determinación



# 1. INTRODUCCIÓN

La minería es una de las actividades económicas más importantes a nivel global, centrada en la explotación y comercialización de minerales metálicos, no metálicos e industriales. Es la base de la economía chilena, ligada al desarrollo histórico del país, cuyo principal producto es el cobre, situando a Chile como el mayor productor del mundo [1]. A medida que la industria minera se ha desarrollado, Chile enfrenta desafíos como: mayor dificultad para acceder al mineral, estrictas regulaciones ambientales y la necesidad de optimizar los insumos críticos para la producción de cobre. Estos puntos son fundamentales para el crecimiento de la industria y deben ir alineados con las transformaciones tecnológicas que la misma ha permitido desarrollar [2]. Frente a estos desafíos es importante considerar esta información para una correcta toma de decisiones en el diseño de planes de producción. Estos últimos son herramientas que permiten: planificar, ejecutar y controlar las actividades en desarrollo y también, un correcto uso de los insumos necesarios para el procesamiento del cobre, tales como: agua, electricidad y las emisiones de gases de efecto invernadero (GEI). Estos son factores claves para la sustentabilidad de la minería y del medio ambiente, y deben ser considerados al realizar una proyección de producción de cobre sustentable.

Los modelos de Redes Neuronales Artificiales (RNA) son apropiados para realizar proyecciones en aplicaciones lineales y no lineales [3], poseen capacidad para: aprender constantemente, entrenarse por sí mismas y mejorar [3]. Las Redes Neuronales Recurrentes (RNR), son un tipo de RNA, y se han aplicado ampliamente para realizar proyecciones en diversas áreas, aun así, son inadecuados a datos de series de tiempo de intervalos largos [4]. Para superar esta limitación, este estudio considera modelos de redes con memoria a corto largo plazo, es decir (LSTM) como una extensión de las RNR y han demostrado una alta eficiencia en este tipo de redes para manejar datos temporales dinámicos [4].

Los autores realizan un modelo basado en técnicas de redes neuronales artificiales del tipo recurrentes para establecer proyecciones de producción de cobre en marco del consumo de agua, energía y emisiones de gases de efecto invernadero y planificar planes de producción de cobre, considerando la sustentabilidad de la minería y el medio ambiente proporcionando una nueva alternativa para los proyectistas de producción en el diseño de sus planes diarios, mensuales y anuales de procesamiento de minerales.

## **1.1. Hipótesis**

Es posible generar un modelo que proyecte la producción de cobre tomando en cuenta el consumo de agua, consumo de energía y emisiones de gases de efecto invernadero mediante Redes Neuronales Recurrentes.

## **1.2. Objetivos**

### **1.2.1. Objetivo general**

Desarrollar un modelo que permita la proyección de producción de cobre, incluyendo factores de consumo de agua, energía y emisiones de gases de efecto invernadero.

### **1.2.2. Objetivos específicos**

- Revisar literatura especializada para elaborar el marco teórico empleado en la construcción del modelo de proyección de producción de cobre.
- Diagnosticar el consumo de agua, energía eléctrica y generación de emisiones en la producción de cobre en Chile.
- Modelar la proyección de producción de cobre en Chile.
- Aplicar metodología de estudio al modelo.
- Generar resultados.
- Realizar análisis y conclusión de resultados obtenidos.

## **1.3. Estructura del informe**

Este trabajo se estructura en el siguiente orden:

- Capítulo 1. Introducción: Se expone el problema y metodología de investigación.
- Capítulo 2. Contexto: Se describe el contexto en que ocurre la investigación.
- Capítulo 3. Marco Teórico: Se describen conceptos para la comprensión de la investigación.
- Capítulo 4. Revisión de la literatura: Se describen estudios relacionados a la investigación.
- Capítulo 5. Metodología de estudio.
- Capítulo 6. Resultados y discusiones.
- Capítulo 7. Conclusiones.

## 2. CONTEXTO

### 2.1. Producción de cobre en el mundo

En los últimos años, la demanda de cobre ha crecido rápidamente como resultado del aumento de la población mundial, el crecimiento económico en las economías emergentes y la transición hacia una sociedad más sostenible [1].

Chile, es el mayor productor y exportador de cobre en el mundo, produjo 5,82 millones de toneladas el año 2019, lo que equivale al 28,4% de la producción mundial. Además de ocupar el segundo lugar en la producción mundial de molibdeno, con 53.541 toneladas que corresponde al 18,4% de esa producción. Los principales países productores de cobre a nivel mundial el 2019 fueron: Chile 28,4%, Perú 11,7%, China 7,8%, los que en conjunto alcanzan el 47,9% de la producción mundial [5]. En la Tabla 2.1 se muestran los mayores productores de cobre del mundo.

**Tabla 2.1:** Producción mundial de cobre en miles de toneladas entre los años 2018-2019 [5].

País	Producción		Participación mundial (%)
	2018	2019	
Mundial	20.422	20.532	-
Chile	5.872	5.822	28,4
Perú	2.440	2.400	11,7
China	1.590	1.600	7,8
Estados Unidos	1.220	1.300	6,3
R.D. Congo	1.230	1.300	6,3
Australia	920	960	4,7
Zambia	854	790	3,9

## 2.2. Producción de cobre en Chile

La producción de cobre se lleva a cabo, principalmente, en las regiones de Tarapacá, Antofagasta y Atacama, donde se concentran casi la totalidad de las reservas conocidas del país, que equivalen al 40% de las reservas del planeta [1]. Existen dos tipos de minerales de cobre que se procesan en el país: los minerales de cobre oxidados y minerales de cobre sulfurados, los cuales requieren de procesos distintos de procesamiento. Para los minerales oxidados se utiliza la vía hidrometalúrgica. Mientras que para los minerales sulfurados se utiliza la flotación, ambos procesos se caracterizan por un gran uso de agua y electricidad, que son insumos críticos para la sustentabilidad de la industria minera [7].

El uso de agua, electricidad y las emisiones de gases de efecto invernadero, como consecuencia del procesamiento de minerales de cobre, son factores claves para la sustentabilidad de la minería y del medio ambiente. Por lo que estos deben ser considerados a la hora de realizar una proyección para la producción de cobre sustentable. En Chile, la minería participa en forma relevante en la economía y el desarrollo sustentable del país, representando en el año 2019 una participación en el PIB del 9,4%, siendo la participación de la minería del cobre en un 8,4%, lo que representa casi el 50% de las exportaciones chilenas [6]. Otros indicadores económicos de la producción de cobre se muestran en la Tabla 2.2 como: producción de cobre, exportación y participación de este en Chile.

**Tabla 2.2.** Indicadores económicos de producción y exportación de cobre entre los años 2010-2019 [6].

Años	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Producción de cobre (Miles de toneladas)	5.457	5.257	5.485	5.851	5.793	5.833	5.626	5.558	5.872	5.822
Exportación cobre (Millones US\$)	40.342	42.664	42.322	39.739	37.969	30.374	27.285	34.662	36.367	33.399
Exportación total país (Millones US\$)	71.028	80.586	78.813	77.368	76.649	63.362	60.597	69.230	75.482	69.686
Participación (US\$) del Cobre (%)	56,8	52,9	53,7	51,4	49,5	47,9	45,0	50,1	48,0	48,0
PIB (%)	16,0	14,9	12,7	11,1	11,2	9	7,8	10,0	9,5	8,4

## **2.3. Insumos críticos para la producción de cobre en Chile**

En base a una adaptación de las definiciones contenidas en [8], se define a un insumo crítico como aquel cuya situación de abastecimiento podría resultar ser crítica para el proyecto u operación en el mediano y largo plazo. A la definición contribuyen principalmente los riesgos de desabastecimiento, alta dependencia de las importaciones o un monopolio de los insumos en pocos oferentes, como también la importancia o peso relativo del insumo en la estructura de costos.

Existen factores que resultan estratégicos para la industria como son los depósitos minerales y las condiciones institucionales para su explotación. Además, otros sobre los que existe un consenso a nivel país, que se deben monitorear como la energía eléctrica y el uso del agua. También existen otros recursos sobre los que se han hecho esfuerzos aislados y no sistemáticos como son el capital humano o algunos más específicos como el ácido sulfúrico o el cianuro [9, 10].

### **2.3.1. Consumo de agua**

El agua es esencial para la vida, las comunidades, el medio ambiente y la actividad económica e industrial [9]. Dado el escenario de escasez hídrica que se ha observado en los últimos años, la competencia por el agua continúa creciendo y el Foro Económico Mundial reconoce una crisis mundial del agua, y la clasifica como uno de los tres primeros riesgos mundiales [9]. Dicha situación es posible revertirla sólo si se toman las medidas necesarias para una gestión más eficiente del recurso hídrico [10].

En este contexto de crisis hídrica a nivel mundial, Chile no está ausente, y según estimaciones de [11]. Su índice de estrés hídrico comparado con los países en más riesgo de una crisis del agua, Chile ocupa el lugar 18, calificado con riesgo alto. Por otra parte, el aumento de la población y el crecimiento de las industrias más importantes del país (agricultura y minería), suponen una mayor demanda de agua. Por ello una adecuada gestión de los recursos hídricos en minería es primordial para un desarrollo sostenible de la industria, y en línea con los objetivos de desarrollo sostenible de la Organización de las Naciones Unidas al año 2030 [12]. Si bien el uso del agua en la industria minera representa solo el 3% del suministro de agua de Chile [13], muchas de las concesiones y operaciones mineras se ubican en zonas donde la escasez de agua es un factor limitante para el desarrollo regional.

Al analizar la situación hídrica, geográficamente en el norte de Chile en donde se concentra la minería, es una de las áreas más secas del planeta y los recursos hídricos superficiales son escasos. Además, existe una demanda creciente de agua por parte de los usuarios industriales, comunidades y el medio ambiente. En este sentido, es importante el esfuerzo realizado por las empresas mineras en apuntar a otras fuentes de extracción como es el agua de mar, con el fin de desestresar las fuentes continentales de agua fresca para sus operaciones.

Considerando el contexto previamente descrito, el suministro de agua es trascendental para la industria minera, debido a la cantidad de procesos que requieren agua para su funcionamiento. Por ello, los impactos en la calidad y cantidad del agua son uno de los aspectos más polémicos de los proyectos mineros.

### **2.3.2. Consumo de electricidad**

La energía eléctrica es un insumo estratégico para la minería del cobre, dado que se requiere en sus diversos procesos productivos y servicios. De acuerdo a estimaciones de [14] a diciembre de 2019, su uso representa alrededor del 9% de los costos operacionales de la gran minería del cobre. Su impacto en el consumo eléctrico del país es significativo, promediando en los últimos 15 años una participación de un tercio en el consumo nacional de energía eléctrica, situación que se puede explicar en gran parte por tres tendencias que han presionado al alza del consumo:

- Caída en las leyes de cobre: Esta situación ha significado que las empresas tengan que extraer grandes volúmenes de mineral para lograr mantener los niveles de producción esperados. Esto ha llevado a un incremento en el uso de energía en procesos como chancado y molienda.
- Creciente uso de agua de mar: Dadas las restricciones para el abastecimiento de agua a través de fuentes continentales, el agua de mar debe ser impulsada desde la costa a las faenas mineras.
- Enfoque en la producción de concentrados de cobre: Proceso que tiene un uso intensivo de energía eléctrica.

Por ello, es relevante dar una mirada prospectiva de su demanda de energía eléctrica al año 2030, en la que podrían estar en operación gran parte de la actual cartera de proyectos de producción de cobre. [2].

### 2.3.3. Emisiones de gases de efecto invernadero

El cambio climático es una problemática que afecta en diferente grado y forma a la mayoría de los países del mundo. El avance de desiertos, períodos de lluvia más prolongados, “*peak*” históricos de calor y frío en diferentes regiones, son algunas de las manifestaciones de este fenómeno. En este sentido, muchas naciones, en especial aquellas de altos ingresos, se encuentran tomando medidas para evitar escenarios adversos de cambio climático determinados por los niveles de emisiones de Gases de Efecto Invernadero (GEI).

Chile, es parte de la Convención y del Protocolo de Kyoto y ha realizado inventarios de sus emisiones de GEI para desarrollar proyectos e iniciativas orientados a su reducción. La minería contribuye a las emisiones de GEI en dos formas principales. En primer lugar, hay emisiones “*in situ*” resultantes de la combustión de combustibles fósiles, asociadas principalmente a los procesos de transporte, calentamiento, secado y procesamiento del material. Por otra parte, y con mayor relevancia, la minería consume grandes cantidades de energía eléctrica, mayoritariamente generada por termoeléctricas, que involucran una alta combustión de combustibles fósiles [15].

Las emisiones de GEI de la minería del cobre crecieron desde el año 2001 en un 95%, mientras la producción de cobre creció sólo un 23,7% [9]. Esto responde, principalmente, a las modificaciones que han experimentado los sistemas de generación eléctrica producto del déficit de abastecimiento de gas natural. Pero también, en menor medida a los incrementos de consumo tanto de combustible, como de energía eléctrica de las áreas mina y tratamiento de minerales. Chile no se encuentra ausente de esta problemática y el año 2008 presentó su Plan de Acción Nacional de Cambio Climático [15] con el objetivo de implementar medidas de orden estratégico y de política pública. De manera de poder reducir las emisiones de GEI proyectadas al año 2030.

En este contexto, la industria minera enfrenta desafíos estratégicos a largo plazo en relación con reducir el consumo de agua, consumo de energía y las emisiones de GEI, debido a las presiones sociales, ambientales y económicas. También el país debe ir en búsqueda de nuevas tecnologías de ingeniería y desarrollo en relación a la reutilización y reciclaje. Junto con una gestión adecuada de las fuentes de suministro y la comprensión de los vínculos entre el agua, energía y emisiones de GEI, para que una mejora de desarrollo en un área, no cree un efecto adverso en otra, sino que la potencie.

## 3. MARCO TEÓRICO

### 3.1. Machine Learning

El *Machine Learning* es una rama de la inteligencia artificial, y esta forma parte de la ciencia de computación, cuyo objetivo es que, los computadores puedan aprender de los datos. De una manera técnica, es el campo de estudio que proporciona a los computadores la habilidad de aprender sin haber sido programados [29]. Por ejemplo, el filtro de los correos spam, es un programa de aprendizaje automático que puede aprender a reconocer mensajes spam, con ejemplos de emails que son spams.

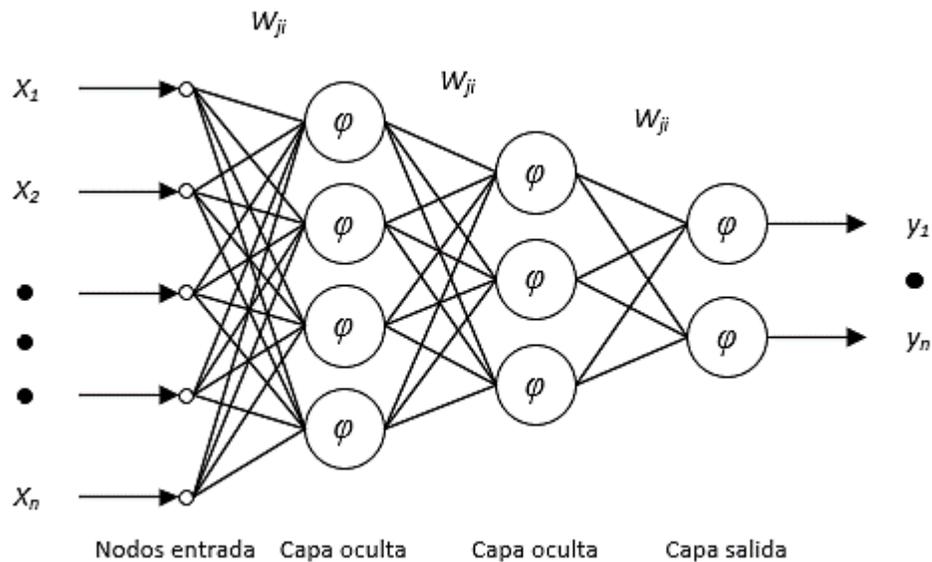
Los campos de *machine learning* y *deep learning*, a veces parecen indistinguibles uno de otro, pero se diferencian en que el *deep learning* eleva el aprendizaje a un nivel más detallado y complejo [29]. Se puede decir que el *machine learning* tiene una vertiente que se denomina *deep learning*. Las dos tecnologías hacen referencia a sistemas capaces de aprender por sí solos, la diferencia entre ambos está en el método de aprendizaje. *Deep learning* es más complejo, sofisticado y más autónomo, lo que quiere decir que, una vez programado el sistema, la intervención del humano es mínima [29].



### 3.2. Redes neuronales artificiales

Las RNA son algoritmos inspirados en el comportamiento de las redes neuronales biológicas existentes en el cerebro. Su objetivo es emular la manera en que dichas redes biológicas: reciben, procesan y generan información. Pueden procesar información en paralelo y en tiempo real, lo que ha permitido su aplicación en la clasificación y reconocimiento de patrones en sistemas complejos [30]. Su principal potencial es, detectar no-linealidades en series temporales, por lo que han sido de gran utilidad en la predicción de datos económicos y financieros [31].

La estructura, se compone de tres capas: la capa de entrada, la capa oculta y la capa de salida [32]. Cada capa tiene un número específico de nodos o neuronas que, interconectan una capa con otra, la conexión entre las neuronas se define por pesos calculados iterativamente en una etapa de entrenamiento [33]. Las redes tienen una inicialización, aleatoria en el proceso de entrenamiento [34], y el aprendizaje, se adquiere en base al ajuste de pesos hasta que, se alcanza un criterio establecido [35], es decir, un criterio para mapear la no linealidad existente entre las entradas y las salidas [36].



**Figura 3.1.** Estructura de una Red Neuronal Artificial [elaboración propia].

En un sistema con RNA, los nodos se conectan por medio de sinapsis, esta estructura de conexión determina, el comportamiento de la red. La Figura 3.1, muestra una estructura de este tipo con dos capas ocultas donde  $X_n$  representan las entradas,  $y_n$  las salidas y  $\varphi$  la función de activación. Las neuronas de cada capa, se interconectan entre si mediante un proceso denominado sinapsis, cuyo valor es diferente para cada una de las conexiones y se representan, mediante los pesos sinápticos  $W_{ji}$ .

La expresión matemática que describe la salida de cada neurona se muestra en la ecuación (1), donde  $W_{ji}$  son los pesos sinápticos que ponderan las entradas  $x_n$ . Cada neurona posee un respectivo nivel de umbral y función de activación  $\varphi$ .

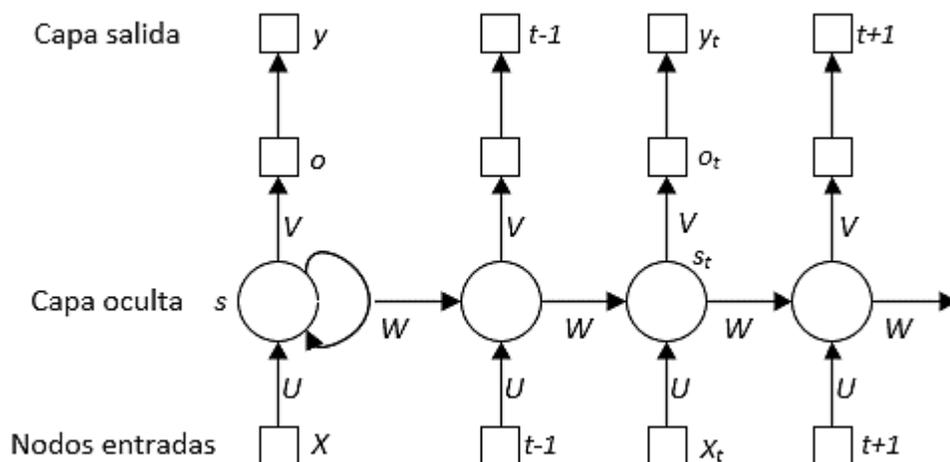
$$y_i = \varphi \left( \sum_{i=0}^n W_{ji} x_n \right) \quad (1)$$

Debido a su constitución y a sus fundamentos, las redes neuronales artificiales presentan un gran número de características semejantes a las del cerebro. Son capaces de aprender de la experiencia, generalizar de casos anteriores a nuevos casos y de separar características esenciales a partir de entradas que representan información irrelevante, entre otros. Esto hace que ofrezcan numerosas ventajas, entre las que se incluyen [37]:

- Aprendizaje adaptativo: Capacidad de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial.
- Autoorganización: Puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje.
- Tolerancia a fallos: A pesar de la destrucción parcial de una red, algunas de sus capacidades se pueden retener, incluso sufriendo un gran daño.
- Operación en tiempo real: Los cálculos neuronales pueden ser realizados en paralelo; para esto se diseñan y fabrican máquinas con *hardwares* especiales para obtener esta capacidad.
- Fácil inserción dentro de la tecnología existente: Se pueden obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilita la integración modular en los sistemas existentes.

### 3.2.1. Redes neuronales recurrentes

En la década de 1980 surgen las primeras RNR [39], el término “recurrente” se refiere a la característica de estas redes de poseer bucles internos de retroalimentación. La RNR convierte los datos de la secuencia de entrada en vectores, y procesa la observación actual en el contexto del estado interno de la red, para generar la salida. La Figura 3.2 muestra la estructura del modelo RNR, en el que  $X_t$  e  $y_t$  son variables de entrada y la variable de salida en el paso  $t$ . El estado oculto  $s_t$  del modelo, se calcula basándose en la entrada  $x_t$  en el paso actual  $t$  y el estado oculto previo  $s_{t-1}$  en el paso  $t-1$ .



**Figura 3.2.** Estructura Red Neuronal Recurrente [elaboración propia].

El modelo matemático de RNR se expresa en las ecuaciones (2), (3) y (4):

$$s^t = f((Ux^t + b) + Ws^{t-1}) \quad (2)$$

$$o^t = Vs^t + c \quad (3)$$

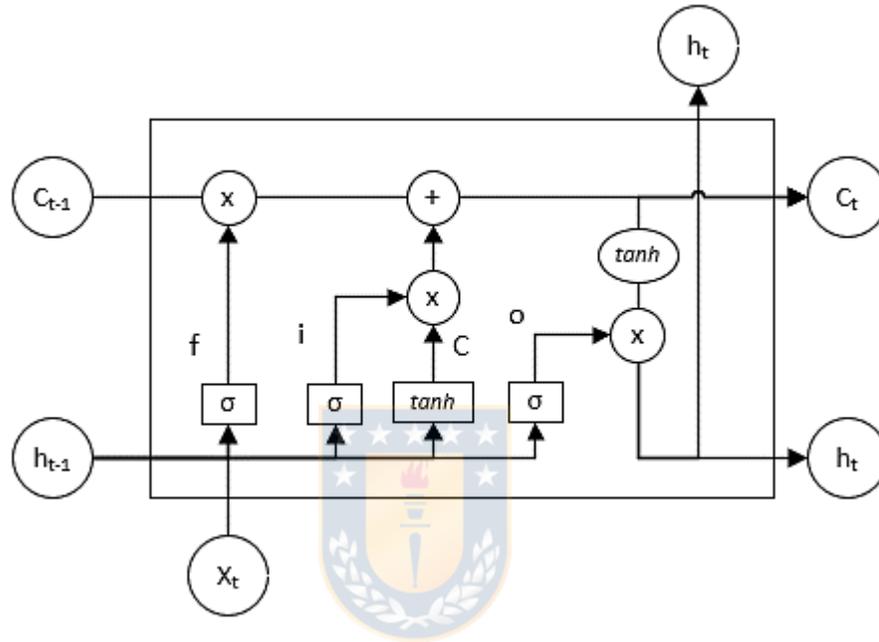
$$y^t = g(o^t) \quad (4)$$

El termino  $U$  es la matriz de peso entre la capa de entrada y la capa oculta. El parámetro  $W$  es la matriz de peso entre las capas ocultas y  $V$ , es la matriz de peso entre la capa oculta y la capa de salida. Se observa en la Figura 3.2, los valores de los parámetros de las matrices de peso  $U$ ,  $W$  y  $V$ , no se modifican en los diferentes pasos. El termino  $s_t$  es el estado oculto, de la capa oculta en el paso  $t$ , este corresponde a la "memoria" de la RNR, la cual determina el termino  $o_t$  que es una variable temporal. Los parámetros  $b$  y  $c$ , son vectores de sesgo y las funciones  $f$  y  $g$  son las funciones de activación de la capa oculta y la capa de salida respectivamente. A pesar de que las RNR son utilizadas para modelar relaciones complejas, estas no son capaces de asimilar dependencias de largo y corto plazo, presentes en datos históricos [38]. Quiere decir, la capacidad de una RNR para identificar y recordar patrones de comportamiento del pasado distante y cercano [39].

Aunque teóricamente las redes RNR deberían retener información en el tiempo  $t$  sobre las entradas, en la práctica, en el largo plazo son imposibles de aprender. Por el estancamiento del descenso del gradiente en un mínimo local de la función, un efecto similar con las RNA que tienen muchas capas de profundidad. A medida que se agregan capas a una red, ésta finalmente se vuelve intransitable [40, 41]. Este error propagado, puede hacer que la disminución del peso del descenso del gradiente, se haga más pequeña y se detenga en las capas inferiores, con el tiempo. Cuando el problema del estancamiento del descenso del gradiente ocurre, puede evitar un correcto entrenamiento de la red [42]. Por lo anterior, Hochreiter, Schmidhuber y Bengio estudiaron las razones teóricas de este efecto a principios de la década de 1990 y surgen en 1997 las RNR de tipo LSTM y GRU [39].

### 3.2.2. Red LSTM

Estas redes poseen bucles de retroalimentación compuestos por células de memoria con un sistema de decisiones basado en compuertas que, se encargan de “recordar” patrones históricos importantes y “olvidar”, patrones no relevantes. Gracias a estas células de memoria, una RNR de tipo LSTM es capaz de asimilar dependencias a corto y largo plazo.



**Figura 3.3.** Estructura Red Neuronal Recurrente tipo LSTM [elaboración propia].

En la Figura 3.3 se presenta la estructura de una red LSTM, este tipo de red consta de tres puertas: una puerta de olvido, una puerta de entrada y una puerta de salida, además, del estado de la celda. La entrada en el paso de tiempo actual, se combina con información sobre el estado oculto en el paso de tiempo anterior y se pasa a través de una función de activación, a la puerta de entrada según, la ecuación (5) [43]:

$$f_t = \sigma(x_t W^f + S_{t-1} U^f + b_f) \quad (5)$$

Donde  $f_t$ , es la puerta de olvido,  $x_t$  es la entrada en el paso de tiempo  $t$  y  $S_{t-1}$  representa el estado oculto en el paso de tiempo anterior  $t-1$ .  $W^f$  es el peso, de la capa de entrada y  $U^f$  es el peso recurrente, del estado oculto. El  $b_f$  es el sesgo, de la capa de entrada. La puerta de entrada tiene dos tareas. La primera, decide qué información se almacena en el estado de la celda para, determinar qué valor de los pesos se actualizan. Mientras que, en la segunda, la capa  $\tanh$  regula la

red, creando un vector de todos los nuevos valores candidatos. Las ecuaciones (6), (7) y (8) representan la ejecución de estas tareas [43, 44]:

$$i_t = \sigma(X_t W^i + S_{t-1} U^i + b_i) \quad (6)$$

$$\hat{C}_t = \tanh(X_t W^c + S_{t-1} U^c + b_c) \quad (7)$$

$$C_t = C_{t-1} f_t + i_t \hat{C}_t \quad (8)$$

Por otro lado, la puerta de salida, decide qué estados ocultos se utilizan para la predicción a través de una función de activación sigmoidea. El nuevo estado de celda modificado, se pasa a la función *tanh* y se multiplica para producir la salida, según las ecuaciones (9) y (10) [44]:

$$O_t = \sigma(X_t W^o + S_{t-1} U^o + b_o) \quad (9)$$

$$S_t = O_t \tanh(C_t) \quad (10)$$

Los pesos, se pueden representar mediante dos matrices,  $W$  y  $U$ . La matriz  $W = \{W^i, W^c, W^o\}$ , representa los pesos de entrada asociados con la capa de entrada, la capa oculta y la capa de salida, respectivamente. Mientras que  $U = \{U^i, U^c, U^o\}$ , representa los pesos del estado oculto anterior, es decir, pesos recurrentes en la capa de entrada, la capa oculta y la capa de salida. Los vectores de sesgo, se pueden entregar como, una matriz  $B = \{b_i, b_c, b_o\}$ , con los índices correspondientes, a las mismas capas descritas anteriormente [43, 44].

### 3.2.3. Red GRU

La red GRU es una nueva generación de RNR introducida, en 2014 por [45]. La red GRU es similar a la red LSTM, pero, tiene una estructura menos compleja y puede, controlar el flujo de información sin la necesidad de utilizar una celda de memoria separada [45]. Las redes GRU, se han aplicado con éxito en tareas de: reconocimiento de voz, mejora del habla, detección de actividad del habla y la recomendación de textos [45]. Una característica distintiva de la red GRU, es su capacidad para capturar dependencias de series de tiempo prolongadas sin requerir celdas de memoria interna, a diferencia de la red LSTM. Esto resuelve el problema del estancamiento en mínimos locales, del descenso del gradiente [45]. Como es una estructura menos compleja, requiere menos recursos computacionales para su construcción y por lo tanto, menores tiempos de trabajo para obtener una solución [45].



$$R_t = \sigma(X_t W^r + S_{t-1} U^r + b_r) \quad (12)$$

$$\hat{h}_t = \tanh(X_t W + U(R_t \sigma(h_{t-1}))) \quad (13)$$

$$h_t = Z_t \sigma(h_{t-1}) + (1 - Z_t) \sigma(\hat{h}_t) + b_h \quad (14)$$

La puerta de actualización, decide qué información se debe recopilar del contenido de la memoria actual y los pasos anteriores. El nuevo estado oculto debe almacenarse en la memoria final, como salida para el paso de tiempo actual y pasar, a través de la función de activación sigmoidea  $\sigma$ .

### 3.3. Series de tiempo

Las series de tiempos, son secuencias de observaciones medidas sobre intervalos de tiempo separados, de manera uniforme y ordenados cronológicamente, donde los datos pueden ser o no dependientes entre sí [46]. Los componentes de una serie de tiempo son:

- Tendencia: La tendencia a largo plazo de una serie creciente o decreciente.
- Estacionalidad: La fluctuación periódica en las series de tiempo dentro de un período determinado forman un patrón que tiende a repetirse de un período estacional al siguiente.
- Ciclos: Largas desviaciones de la tendencia debido a factores diferentes de la estacionalidad durante un intervalo de tiempo extenso.
- Movimiento irregular: Ruido aleatorio o error en una serie de tiempo.

Por esta razón, el pronóstico de series de tiempo se considera, como uno de los diez problemas más desafiantes en la minería de datos, debido a sus propiedades únicas [47]. Actualmente, las series de tiempo juegan un papel imperativo en varios problemas del mundo real, como los mercados financieros, el tráfico de la red, la previsión meteorológica y la producción industrial, entre otros [46]. En el pasado, el problema de series de tiempo ha sido tratado por métodos lineales para lograr las actividades de pronóstico. Sin embargo, en las últimas dos décadas, varios algoritmos de RNA se han establecido como serios contendientes a los métodos estadísticos después de mostrar mejores precisiones de predicción [48]. Dados los diversos algoritmos de RNA, la identificación de un algoritmo específico para una tarea de pronóstico debe basarse en un compromiso entre tres aspectos a saber: la complejidad de la solución, la precisión de la predicción deseada y las características de los datos [48]. Este artículo, se enfoca en la predicción de la producción de cobre, mediante distintos tipos de RNR.

## 4. REVISIÓN DE LA LITERATURA

En este capítulo se revisa la literatura relevante relacionada a las RNA y su aporte en proyecciones de series de tiempo. Primero, se revisan estudios relacionados con el diseño de RNA para proyectar la producción de diversos recursos. En segundo lugar, se analiza la literatura disponible sobre RNR para proyectar series de tiempo, con énfasis en las investigaciones de los últimos años.

### 4.1. Redes neuronales y las proyecciones de producción

En el artículo [16], se presentan dos modelos de redes neuronales para predecir la producción de cobre, mediante el análisis de insumos críticos para la producción, como el consumo de electricidad y el volumen de producción de ácido sulfúrico. El primer modelo considera parámetros de consumo de electricidad para predecir el volumen de producción de cobre. El segundo modelo incluye un parámetro adicional, el volumen de producción de ácido sulfúrico. En el estudio [17] se propone, un método basado en una red neuronal multicapa (RNM) para predecir la producción de mineral mediante, sistemas de transporte de camiones en minas a cielo abierto. El método propuesto utiliza dos modelos RNM diseñados para predecir la producción durante las sesiones de transporte de la mañana y la tarde, respectivamente. La configuración de los nodos de entrada de los modelos RNM se basa en las condiciones de consumo de electricidad y los tiempos de operación.

En el estudio [18] proponen una RNA para modelar las tasas de producción de: petróleo, gas y agua, en función de: presión, inyección y producción. El modelo de predicción de producción para el embalse de inundación de agua se construye, entrenando distintas arquitecturas de la red neuronal con distintos números de capas y de neuronas. Utilizan como métodos de entrenamiento los algoritmos de: Levenberg-Marquardt, Gradiente conjugado escalado y Regularización Bayesiana. En el estudio [19] proponen, una RNM para predecir la producción de petróleo a largo plazo. Consideran tres modelos, el primero es univariante para la predicción. El segundo es multivariable con distintas configuraciones generalizando el primero. Mientras, el tercero es un modelo multivariable de múltiples pasos. Estos modelos se entrenan para predecir repetidamente varios pasos en la serie de tiempo hasta alcanzar el horizonte deseado en un rango de confianza establecido.

En el estudio [20] se desarrolló una RNA que modela la producción de metano. Se pronostica la acumulación de metano en función de: la temperatura de operación, la composición del sustrato y la dosis química. En el artículo [21], proponen un enfoque de RNA clásico y un modelo RNA integrado, con análisis envolvente de datos (AED) (RNA-AED) para la predicción de producción de etileno. El enfoque propuesto utiliza, el modelo AED con variables de holgura para realizar un análisis de sensibilidad. De este modo, determinar las unidades de toma de decisiones efectivas e indicar la dirección optimizada en función de las decisiones ineficaces. En comparación con el enfoque RNA tradicional, el modelo de predicción RNA-AED verifica, de manera efectiva, mediante la ejecución de una comparación lineal entre todas las decisiones consideradas y las decisiones efectivas.

En el artículo de [22] usan: algoritmos de aprendizaje profundo, máquina de vectores de soporte y RNA para pronosticar las emisiones de GEI. El rendimiento de los algoritmos, se analiza con las métricas (MSE, MAE, MAPE y  $R^2$ ), además, como variables para la predicción consideran, el consumo de electricidad y combustibles en la producción. En el estudio [23] aplicaron un modelo RNA con algoritmo de colonia de abejas (ACA) para estimar la producción anual de energía eléctrica en Turquía. Las variables independientes del modelo son: demanda bruta de energía eléctrica, población, temperatura media anual y consumo de energía. El modelo fue entrenado con el algoritmo de retropropagación y para, evaluar su precisión se aplicaron las métricas MSE y  $R^2$ .

#### **4.2. Redes neuronales recurrentes y las proyecciones de producción**

En el estudio [24] proponen, un modelo de pronóstico para la producción de petróleo, cuyo modelo es una RNR con puertas profundas múltiples (RNRM). El modelo propuesto dispone de una arquitectura de baja complejidad y la capacidad, de rastrear conjuntos de datos de series de tiempo de intervalo amplio. Para evaluar el modelo, la técnica propuesta se comparó con una RNM y un modelo LSTM. El estudio [25], propone un modelo LSTM para la predicción de producción de petróleo, basado en la descomposición de los datos de modo empírico por conjuntos para, la predicción de la producción. Los datos del conjunto de prueba se agregan gradualmente al conjunto de entrenamiento y se descomponen para, obtener múltiples funciones de modo intrínseco. La estabilidad de las funciones se analiza por medias y la similitud de las curvas calculadas, luego se selecciona el número de funciones estables como variables predictoras para el aprendizaje automático. Los hiperparámetros óptimos de LSTM son determinados con un algoritmo genético.

Por otro lado, en el estudio [26] los autores presentan, un modelo LSTM para pronosticar la producción de petróleo. Se aplica un algoritmo genético para, configurar de manera óptima la arquitectura y los parámetros de la red LSTM. Al mismo tiempo, comparan el rendimiento de la LSTM con el rendimiento de otros modelos de referencia utilizando, los mismos conjuntos de datos y las mismas condiciones experimentales mediante diferentes medidas de error. Los modelos de referencia varían desde: métodos estadísticos, métodos de RNA y métodos híbridos.

En el estudio [27] proponen, pronosticar el consumo de etanol en Brasil con un enfoque de modelos LSTM y Promedio Móvil Integrado Autorregresivo (ARIMA) considerando, modelos univariados y multivariados para cada caso. Finalmente, en el estudio [28] se propone un método de predicción basado en una RNR para pronosticar la producción de energía fotovoltaica. Los autores utilizan como variables independientes: días adyacentes y datos intradiarios, para mitigar los efectos de las características de no linealidad que existen entre la potencia de salida fotovoltaica y la precisión de la predicción. El modelo es comparado: con una red neuronal de retropropagación, una red neuronal de función de base radial y la máquina de vectores de soporte como puntos de referencia.

La revisión realizada, coloca en contexto los avances en la literatura, sobre el uso de las metodologías de redes neuronales en la confección de modelos, para realizar proyecciones de producción en diferentes áreas. Este estudio, es el primero en analizar la producción de cobre mediante redes neuronales. En la Tabla 4.1, se presenta un sumario de los artículos revisados en esta sección, señalando: Autores, Área de aplicación, Métodos de solución, Variables utilizadas y *Software* utilizado.

### 4.3. Brecha del conocimiento

**Tabla 4.1.** Resumen de estudios disponibles en la literatura especializada

Autor(es) / Cita	Área de aplicación	Metodología	Variables Independientes							Software
			Producción previa	Consumo electricidad	Consumo agua	Consumo petróleo	Precio	Consumo ácido sulfúrico	Temperatura	
Ivana Veličkovska et al (2020)	Producción de cobre	RNA		X					X	SPSS
Jieun Baek & Yosoon Choi (2020)	Producción mineral	RNAM	X	X						Arena
Igor Aizenberg et al (2016)	Producción de petróleo	RNAM	X							
Fares Almomani et al (2020)	Producción de metano	RNA	X						X	TLAB
Yong-Ming Han et al (2016)	Producción de etileno	RNA-AED	X		X	X				
Negash Mamo & Yaw Dennis (2020)	Producción de hidrocarburos	RNA RNAM	X		X					
Melahat Bakay & Ümit Agbulut (2021)	Producción de gases de efecto invernadero (GEI)	RNA RNAM		X		X				
Ergun Uzlu et al (2014)	Producción de electricidad	RNA-ACA		X					X	
Raghad Al-Shabandar et al (2021)	Producción de petróleo	RNR GRU	X	X						SaaS
J. A. Puentes et al (2021)	Producción de etanol	LSTM ARIMA					X		X	SSD SATA
Wei Liu et al (2020)	Producción de petróleo	LSTM	X							TensorFlow
Gangqiang Li et al (2019)	Producción de electricidad	RNR	X			X				
Alaa Sagheera & Mostafa Kotbb (2019)	Producción de petróleo	LSTM	X	X			X			Statsmodels
<b>P. Contreras et al (2021)</b>	<b>Producción de Cobre</b>	<b>RNR GRU LSTM</b>	<b>X</b>	<b>X</b>	<b>X</b>		<b>X</b>			<b>Python</b>

## 5. METODOLOGÍA DE ESTUDIO

En el presente estudio, se utilizaron tres modelos de RNR, el primero, es una red neuronal recurrente simple, el segundo un modelo de red GRU y el tercero un modelo de red LSTM. Todos los modelos consideran información oficial, disponible en el Ministerio de Minería y el Banco Central de Chile.

### 5.1. Conjunto de datos

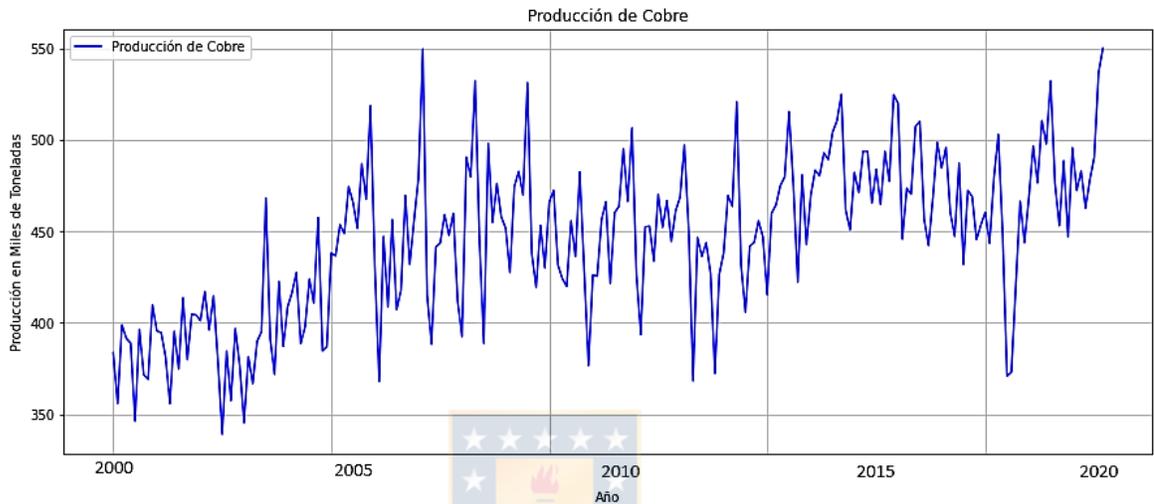
El conjunto de datos de distribución, insumos energéticos, emisiones y producción de cobre de Chile está disponible en el sitio web del Ministerio de Minería. En el conjunto de datos se consideran, cinco conjuntos de características, tales como: producción de cobre, consumo de agua, consumo de energía eléctrica, precio de la libra de cobre y emisiones de GEI. En la Tabla 5.1, se presenta el conjunto de datos, que contiene 1.500 observaciones de producción de cobre en Chile, durante un período entre los años 2000-2020.

*Tabla 5.1. Matriz de datos para producción de cobre.*

Fecha	Producción Cu (Miles Ton)	Precio (cUS\$/libra)	Consumo de agua (L/s)	Consumo de electricidad (TJ)	Emisiones GEI (Millones KT)
01-01-2000	383,6	131,8	1228,7	2689,4	230,8
01-02-2000	355,8	127,2	1316,6	2606,1	230,7
01-03-2000	398,8	121,9	1208,7	2514,1	209,7
01-04-2000	391,5	117,8	1236,5	2464,6	199,1
01-05-2000	388,9	124,4	1091,3	2441,9	207,6
...	...	...	...	...	...
01-06-2020	466,5	275,5	1068,3	7538,1	415,3
01-07-2020	464,8	302,5	1063,3	7617,7	424,7
01-08-2020	484,2	307,1	1080,8	7519,3	425,9
01-09-2020	479,8	317,3	1073,3	7491,5	425,1
01-10-2020	486,5	319,6	1067,5	7531,6	413,9
01-11-2020	485,1	335,4	1071,6	7511,6	430,8
01-12-2020	499,7	366,7	1051,6	7361,1	428,5

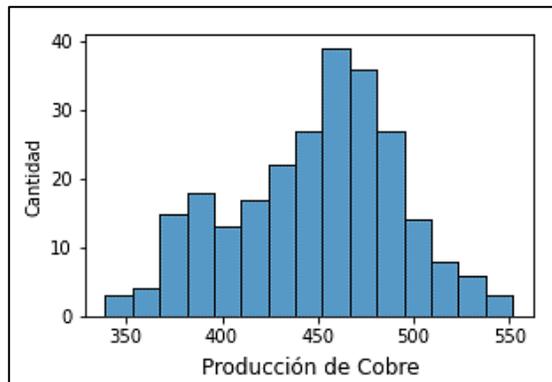
### 5.1.1. Producción de Cobre

La Figura 5.1, se utilizó para visualizar la diferencia en la producción de cobre a nivel nacional a lo largo del tiempo. Se observa que, existen fluctuaciones aleatorias en los datos, particularmente, en el período comprendido entre los años 2000 y 2005. Sin embargo, se muestra una variación constante, en la tasa de producción de cobre, durante los últimos 10 años.



**Figura 5.1.** Producción de cobre mensual en Chile, años 2000-2020 [elaboración propia].

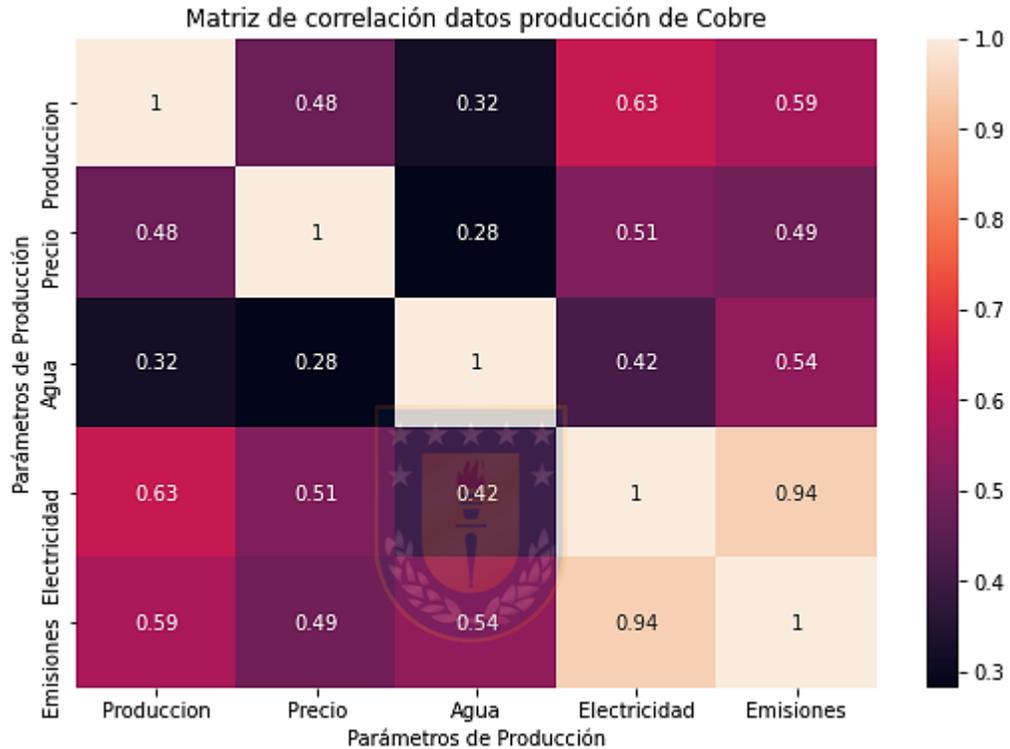
Se puede observar que el modelo lineal no se ajusta a los datos, y como consecuencia, se debe considerar un método de pronóstico más avanzado. El histograma de la Figura 5.2 muestra que, la distribución de los datos no sigue distribución normal y sesgada a la derecha de la curva. Los datos fueron normalizados evitando, ingresar a la red neuronal valores elevados provocando que los modelos propuestos, tengan un rendimiento deficiente [49].



**Figura 5.2.** Histograma producción de cobre en Chile [elaboración propia].

### 5.1.2. Matriz de correlación datos de producción de cobre

Se calcula el valor del coeficiente de correlación de Pearson entre las variables independientes y la producción de cobre. En la Figura 5.3, se observa que los datos están relacionados unos con otros, sobre todo en el consumo de electricidad con las emisiones de GEI. Por otro lado, existe una baja correlación entre el consumo de agua y la producción.

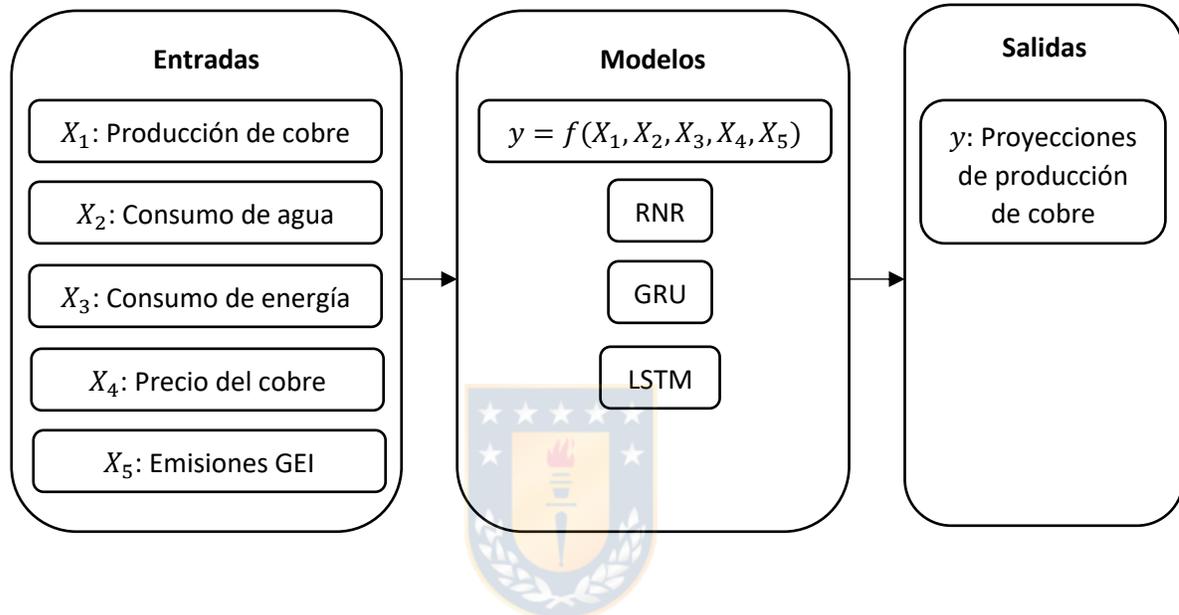


**Figura 5.3.** Matriz de correlación datos producción de cobre [elaboración propia].

El conjunto de datos se dividió en dos subconjuntos, es decir, los conjuntos de entrenamiento (75%) y validación (25%). A su vez los datos de entrenamiento se dividieron en pequeños lotes de muestras de diferentes tamaños, y luego son ingresados al modelo. Dado que el objetivo de este estudio es pronosticar, la producción mensual de cobre, se establece, el paso de tiempo en "1". Este estudio, utilizó tres modelos de redes neuronales, incluidas las redes tipo RNN, GRU y LSTM, con diferentes valores de hiperparámetros. Estos métodos se proyectaron para los próximos 12 meses desde la última fecha, el dato resultante es la cantidad de cobre producido, en los próximos meses. Para el estudio, se ejecutaron los modelos utilizando la biblioteca de Python, Keras.

## 5.2. Modelos

Se realizaron predicciones, para tres modelos de RNA, redes tipo RNR, GRU y LSTM. En cada modelo se consideró cinco variables independientes de entrada  $X_i$ , para modelar la producción de cobre mensual, en miles de toneladas de cobre. En la Figura 5.4 se muestra el esquema general de los modelos y se detallan las variables de entrada.



**Figura 5.4.** Metodología para los modelos RNR, GRU y LSTM [elaboración propia].

El modelo considera cinco características de entrada denotadas por  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$  y  $X_5$ .  $X_1$ , es la producción de cobre mensual;  $X_2$ , es el consumo de agua mensual durante la producción de cobre,  $X_3$ , es el consumo de energía eléctrica mensual durante la producción de cobre;  $X_4$ , es el precio del cobre en centavos por libra y  $X_5$ , corresponde a las emisiones de GEI durante la producción de cobre. La variable  $y$  corresponde, al tonelaje de cobre producido mensualmente.

El número de parámetros de entrada y el número de neuronas ocultas de los modelos de RNA, tienen un impacto significativo en el resultado final de la predicción del modelo, sin embargo, existe una investigación limitada sobre la arquitectura óptima de los modelos RNA. En este estudio, se establecieron diferentes números de neuronas ocultas, en la construcción del modelo RNA y se determina, la arquitectura óptima de éste, en función de los resultados del rendimiento de entrenamiento y validación. Durante la selección del número de neuronas, es importante evitar

excesos de parámetros y el sobreajuste del modelo ya que, resultaría en fallas durante la predicción. Porque, los modelos sobreajustados extraerían parte de la variación residual y se estancarían, en mínimos locales. Para evitar el sobreajuste, el número de capas ocultas se mantiene al mínimo tanto, como sea posible. Se realiza, una serie de pruebas sobre el entrenamiento y la validación, del conjunto de datos para determinar la arquitectura óptima de los modelos.

Los modelos entrenados, se utilizan para generar la salida prevista de la predicción y realizar una comparación con los valores de salida, del conjunto de datos de validación. En relación a métricas de evaluación para los resultados de la regresión, las utilizadas en el presente estudio, son las siguientes:

- **Coefficiente de determinación  $R^2$** : Indica la calidad del modelo frente a los resultados, y es definido por la ecuación (15).

$$R^2 = \frac{\sigma_{XY}^2}{\sigma_X^2 \sigma_Y^2} \quad (15)$$

Con  $\sigma_{XY}$ , covarianza de  $(X, Y)$ ,  $\sigma_X^2$ , la varianza de  $X$ ,  $\sigma_Y^2$ , la varianza de  $Y$ .

- **Mean Absolute Error (MAE)**: Es la diferencia absoluta entre el valor objetivo y el valor predicho, se calcula según la ecuación (16).

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (16)$$

- **Mean Square Error (MSE)**: Corresponde al promedio de la diferencia cuadrada entre el valor objetivo y el valor predicho, se calcula según la ecuación (17).

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (17)$$

- **Mean Absolute Percentage Error (MAPE)**: Versión porcentual de MAE, entrega la relación entre que tan lejos están las predicciones obtenidas con el valor real promedio, se calcula según la ecuación (18).

$$MAPE = \frac{100}{N} \sum \left[ \frac{|y_i - \hat{y}_i|}{y_i} \right] \quad (18)$$

Para los modelos se utilizaron como hiperparametros generales: la función de perdida MSE durante el entrenamiento; funciones de activación tangente hiperbólica y sigmoidea; función de optimización RMSprop y *batch size* de 12. En Tabla 5.2, se muestran las configuraciones de los modelos de redes neuronales a usadas.

**Tabla 5.2.** *Parámetros de configuración experimentales de los modelos RNR, GRU y LSTM.*

N° Capas ocultas	N° Neuronas por capa	Tipo de Red
1	16	RNR
	32	GRU
	64	LSTM
	96	
2	16-16	RNR
	16-32	GRU
	32-32	LSTM
	32-64	
	64-64	
	64-96	
3	96-96	
	16-32-32	RNR
	32-64-64	GRU
	64-96-96	LSTM

Los modelos se ejecutaron en un computador Asus, Zenbook Pro, Intel core i9, CPU 5.3 GHz, 32 GB de memoria RAM, en lenguaje Python mediante el programa Jupyter Notebook.

## 6. RESULTADOS

En esta sección se presentan los resultados de la simulación para los modelos de producción de cobre. El análisis considera la predicción para 60 meses de producción de cobre con los modelos RNR, GRU y LSTM. Dado que el tamaño del lote de datos juega un papel importante para evitar que el proceso de aprendizaje se detenga en un mínimo local, los conjuntos de entrenamiento y validación se dividieron en 12 lotes. Se evalúan las métricas MAE, MAPE, MSE y  $R^2$  para los conjuntos de entrenamiento y validación, destacando en negrita los mejores resultados.  $R^2$  y MAPE no se asocian a unidades de medida, un mayor  $R^2$  y menor valor del MAPE indican, un mejor rendimiento del modelo. Las unidades de MSE y MAE son miles de toneladas de cobre fino.

### 6.1. Resultados modelo RNR

Se observa que distintas arquitecturas de redes neuronales generan diferentes errores. Los resultados obtenidos para las arquitecturas evaluadas, se presentan en la Tabla 6.1. Las primeras dos columnas corresponden al número de capas y al número de neuronas respectivamente, y las columnas restantes contienen los valores de MAE; MAPE, MSE y  $R^2$ . Para un mismo número de capas, se observa, que a medida que aumenta el número de neuronas, se obtiene una disminución del error de predicción. Así, en la red neuronal con dos capas ocultas, con 64 neuronas por capa, produce un  $R^2 = 0,25$ . Además, MAE disminuye de 36,58 a 21,97. En general el aumento en el número de neuronas y capas genera una disminución también para MAPE y MSE.

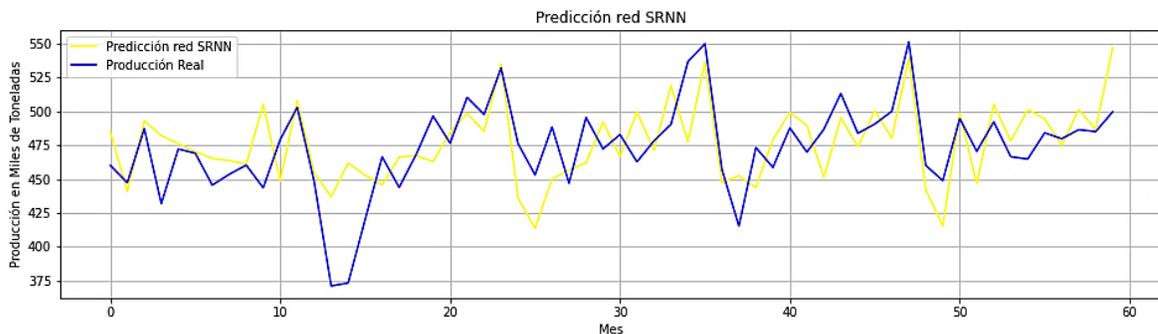
La disminución también se constata en las redes con una y tres capas. Sin embargo, en la red con dos capas la disminución ocurre solamente hasta la red con 128 neuronas, que genera el mínimo error MAE = 22,07. Este hecho sugiere que es un tamaño cercano al óptimo, pues las redes de mayor tamaño generan un aumento del error MAE = 30,23.

En la Figura 6.1 se presenta la predicción realizada con la red de dos capas ocultas, con 64 neuronas por capa. El gráfico presenta la predicción y la producción real de cobre en cada mes. Este se realiza utilizando 60 datos de producción real y 60 datos predichos por el modelo. Los mejores resultados son reflejados gráficamente cuando, los valores predichos son más cercanos a los valores reales.

**Tabla 6.1.** Resultados de métricas: MAE, MAPE, MSE y  $R^2$  en predicción de producción de cobre con modelo RNR.

N° Capas	N° Neuronas	MAE	MAPE	MSE	$R^2$
1	16	36,58	0,0774	2077,75	-0,91
	32	31,73	0,0678	1537,78	-0,42
	64	30,52	0,0657	1490,49	-0,37
	96	26,96	0,0578	1151,61	-0,06
2	16-16	31,25	0,0618	1419,09	-0,31
	16-32	29,07	0,0576	1279,01	-0,18
	32-32	26,42	0,0522	1135,26	-0,05
	32-64	23,67	0,0482	997,79	0,08
	<b>64-64</b>	<b>22,07</b>	<b>0,0522</b>	<b>810,81</b>	<b>0,25</b>
	64-96	23,86	0,0511	973,60	0,10
	96-96	23,35	0,0565	1089,78	0,00
3	16-32-32	30,23	0,0664	1588,58	-0,46
	32-64-64	26,20	0,0565	1055,23	0,03
	<b>64-96-96</b>	<b>21,97</b>	<b>0,0480</b>	<b>846,59</b>	<b>0,22</b>

Considerando producciones entre las 420 y 520 mil toneladas mensuales, los errores indican que se podrían predecir valores de producción coherentes y apropiados, a la producción real. Se observa un error entre el 4.0 y el 7.0% para la red. Este hecho sugiere que, se puede realizar la predicción de la producción mensual de cobre, a partir de datos del pasado.



**Figura 6.1.** Gráfico de predicción de producción de cobre del modelo RNR (64-64) vs producción real.

## 6.2. Resultados modelo GRU

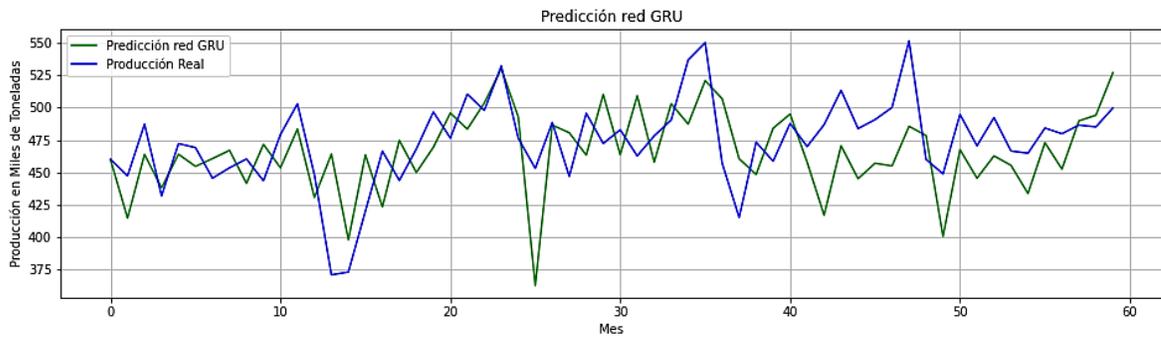
Se observa que distintas arquitecturas de redes generan diferentes errores. Los resultados obtenidos para las arquitecturas evaluadas se presentan en la Tabla 6.2. Al igual que en el modelo RNR, para un mismo número de capas, se observa, que a medida que aumenta el número de neuronas, se obtiene una disminución del error. Así, en la red neuronal con dos capas ocultas, con 64 neuronas por capa, produce un  $R^2 = 0,50$ . Además, MAE disminuye de 38,09 a 26,97. En general el aumento en el número de neuronas y capas genera una disminución también para MAPE y MSE.

La disminución también se constata en las redes con una y tres capas. Sin embargo, en la red con dos capas la disminución ocurre solamente hasta la red con 128 neuronas, que genera el mínimo error MAE = 26,97. Este hecho sugiere que es un tamaño cercano al óptimo, pues las redes de mayor tamaño generan un aumento del error MAE = 30,47.

En la Figura 6.2 se presenta la predicción realizada con la red de dos capas ocultas, con 64 neuronas por capa. El gráfico presenta la predicción y la producción real de cobre en cada mes, descrito anteriormente. Considerando producciones entre las 420 y 520 mil toneladas mensuales, los errores indican que se podrían predecir valores de producción coherentes y apropiados, a la producción real. Se observa un error entre el 5.0 y el 7.0% para la red. Este hecho sugiere que, se puede realizar la predicción de la producción mensual de cobre, a partir de datos del pasado.

**Tabla 6.2.** Resultados de métricas: MAE, MAPE, MSE y  $R^2$  en predicción de producción de cobre con modelo GRU.

N° Capas	N° Neuronas	MAE	MAPE	MSE	$R^2$
1	16	29,72	0,0578	1231,11	-0,13
	32	29,12	0,0631	1331,88	-0,23
	64	29,04	0,0619	1309,65	-0,21
	96	25,22	0,0536	941,83	0,13
2	16-16	33,38	0,0711	1761,84	0,07
	16-32	26,94	0,0584	1158,25	0,38
	32-32	38,09	0,0808	1954,09	0,20
	32-64	34,72	0,0741	1837,59	0,31
	<b>64-64</b>	<b>26,97</b>	<b>0,0579</b>	<b>1045,16</b>	<b>0,50</b>
	64-96	30,47	0,0647	1322,58	0,22
	96-96	27,96	0,0600	1148,14	0,06
3	16-32-32	27,54	0,0563	1164,20	0,02
	32-64-64	27,10	0,0576	1064,82	0,12
	64-96-96	28,84	0,0617	1280,05	0,18



**Figura 6.2.** Gráfico de predicción de producción de cobre del modelo GRU (64-64) vs producción real.

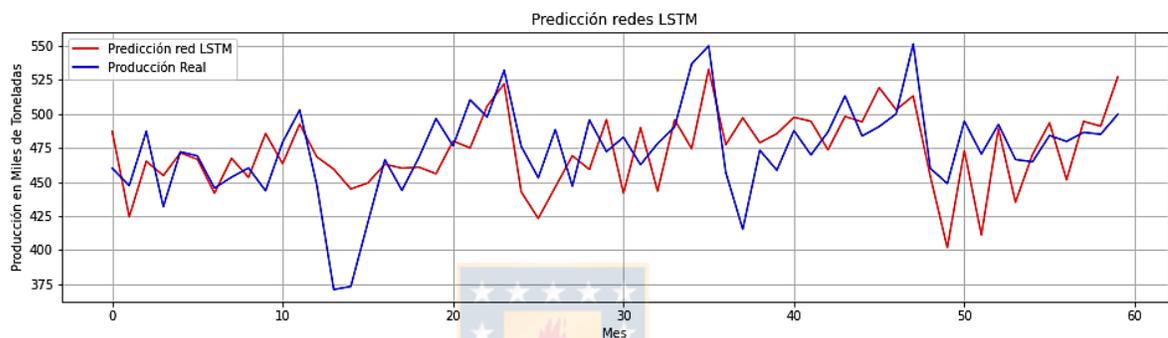
### 6.3. Resultados modelo LSTM

Se observa que distintas arquitecturas de redes generan diferentes errores. Los resultados obtenidos para las arquitecturas evaluadas se presentan en la Tabla 6.3. Al igual que en el modelo RNR y GRU, para un mismo número de capas, se observa, que a medida que aumenta el número de neuronas, se obtiene una disminución del error de predicción. Así, en la red neuronal con dos capas ocultas, con 64 neuronas por capa y en la red con dos capas ocultas, con 64 y 96 neuronas por capa, produce un  $R^2 = 0,76$  y  $R^2 = 0,74$  respectivamente. Además, MAE disminuye de 37,47 a 22,04. En general el aumento en el número de neuronas y capas genera una disminución también para MAPE y MSE.

**Tabla 6.3.** Resultados de métricas: MAE, MAPE, MSE y  $R^2$  en predicción de producción de cobre con modelo LSTM.

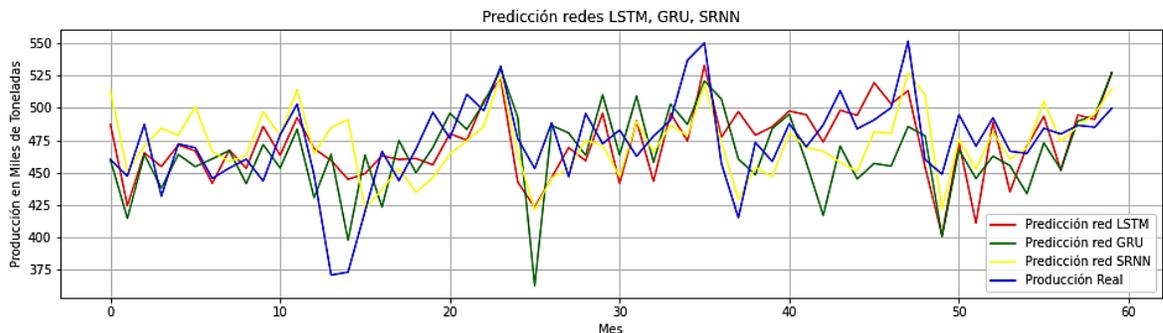
Nº Capas	Nº Neuronas	MAE	MAPE	MSE	$R^2$
1	16	32,98	0,0689	1598,64	-0,47
	32	26,88	0,0569	1127,64	-0,04
	64	25,14	0,0534	933,44	0,14
	96	22,01	0,0475	767,02	0,29
2	16-16	32,47	0,0690	1770,77	-0,63
	16-32	24,02	0,0520	986,70	0,09
	32-32	37,47	0,0786	1896,67	0,13
	32-64	24,63	0,0526	903,69	0,17
	<b>64-64</b>	<b>22,04</b>	<b>0,0277</b>	<b>755,87</b>	<b>0,76</b>
	<b>64-96</b>	<b>22,79</b>	<b>0,0289</b>	<b>851,31</b>	<b>0,74</b>
3	96-96	23,87	0,0520	946,14	0,64
	16-32-32	26,29	0,0559	1147,64	0,06
	32-64-64	21,47	0,0464	857,74	0,65
	64-96-96	11,88	0,0287	800,69	0,69

La disminución se constata en las redes con una y tres capas. Sin embargo, en las redes con dos capas la disminución ocurre solamente hasta las redes con 128 y 160 neuronas respectivamente, que generan el mínimo error MAE = 22,04 y MAE = 22,79. Este hecho sugiere que es un tamaño cercano al óptimo, pues las redes de mayor tamaño generan un aumento del error MAE = 26,29. En la Figura 6.3 se presenta la predicción realizada con la red de dos capas ocultas, con 64 neuronas por capa. Considerando producciones entre las 420 y 520 mil toneladas mensuales, se observa un error entre el 2.0 y el 7.0% para la red. Este hecho sugiere que, se puede realizar la predicción coherente de la producción mensual de cobre, a partir de datos del pasado.



**Figura 6.3.** Gráfico de predicción de producción de cobre del modelo LSTM (64-64) vs producción real.

En la Figura 6.4, se presenta una comparativa de los resultados de los tres modelos en detalle frente a la producción real. Se observa que el modelo LSTM representa un mejor ajuste a los datos de producción, respecto a los demás modelos. El modelo GRU presenta un menor ajuste a los datos respecto al LSTM, pero mejores que los del modelo RNR. Los tiempos de ejecución para todos los modelos fue de 15 segundos.



**Figura 6.4.** Grafica predicción modelos RNR, GRU y LSTM vs producción real.

## 6.4. Discusión

Para el modelo de RNR, los resultados de la simulación indicaron que, los valores  $R^2$  fueron, los más bajos de los modelos propuestos y obtuvieron, los errores más altos de las demás métricas. Esto es de esperar ya que, la estructura de la red es simple respecto a las redes GRU y LSTM que disponen de una unidad de memoria que, permite un mejor rendimiento en el corto y largo plazo. En relación al número de capas y neuronas por capa, se demostró que a medida que estas aumentaban los valores de  $R^2$  aumentaban y las métricas de MAE, MAPE y MSE disminuían. Los investigadores demuestran, que aumentar la cantidad de capas ocultas podría mejorar significativamente, el rendimiento de la red al igual que el número de neuronas [50].

Para el modelo de red neuronal GRU las métricas indican un valor de  $R^2 = 0,50$  para dos capas. Esta mejora respecto a la red RNR es de esperar ya que las redes GRU han demostrado buenos rendimientos al ser utilizadas para resolver problemas de series de tiempo. Específicamente para tareas de pronóstico a largo plazo, ya que puede entrenar conjuntos de datos de intervalo largo, más rápido que otros métodos estándar [24]. Respecto al número de capas y neuronas, los mejores resultados se obtuvieron en los modelos compuestos por dos capas ocultas. Al aumentar el número de capas, el rendimiento mejora hasta cierto punto, en el que, al continuar aumentando su número, el rendimiento desciende. Esto debido a que el aumento de conexiones entre las redes podría generar un sobre entrenamiento de la red y por lo tanto la acumulación de los errores para el descenso del gradiente conjugado [37].

Para el modelo de red neuronal LSTM las métricas indicaron los mejores resultados para el conjunto de datos de validación, ya que presento un valor de  $R^2 = 0,76$  para dos capas. El modelo LSTM mejora respecto al modelo GRU debido a, su unidad de memoria que permite recordar los datos “clave” y olvidar los datos que no son relevantes. Respecto al número de capas y neuronas por capa, los mejores modelos se obtuvieron al igual que para los modelos GRU, con dos capas ocultas. Demostrando que a medida que estas aumentan los valores de  $R^2$  mejoran y las métricas de MAE, MAPE y MSE disminuyen. De igual manera el aumento de capas y neuronas después de este punto genera un aumento en las métricas de error y disminución del término  $R^2$ , como era de esperarse producto de un posible sobre entrenamiento de la red.

La red LSTM tiene una estructura compleja en comparación con GRU, y toda la secuencia de datos debe almacenarse en la memoria. Sin embargo, apilar más capas recurrentes en el modelo propuesto, puede ser una forma de mejorar la representación del modelo, aunque los costos computacionales para esto, podrían ser altos. El uso de capas GRU en un modelo de aprendizaje profundo, proporciona una reducción significativa en el costo del modelo [29]. Con respecto a los datos utilizados, se trabajó con datos informados por la base de datos de COCHILCO en el Ministerio de Minería, que comprende información de las empresas de la gran minería pertenecientes al consejo minero. Además, indicadores económicos informados por el Banco Central.

El escalamiento de los datos pudo haber influido en el comportamiento de la red. Por ejemplo, muchos elementos usados en la función objetivo de un algoritmo de aprendizaje, asumen que todas las características están centradas alrededor de cero y tienen varianzas en el mismo orden. Si una característica tiene una varianza que es, órdenes de magnitud mayor que otras, podría dominar la función objetivo y hacer que el escalador, no pueda aprender de otras características correctamente [29]. En general, los algoritmos de aprendizaje se benefician de la estandarización del conjunto de datos y los comportamientos de los diferentes escaladores, en un conjunto de datos que contiene valores atípicos marginales. Podrían comportarse mal si las características individuales no se parecen a datos estándar distribuidos normalmente: Gaussiano con media cero y varianza unitaria.

Se debe destacar que, los datos pueden complementarse con otros, como la producción de ácido sulfúrico, toneladas de relaves producido y programación de mantenciones de equipos y maquinarias, entre otros. Sin embargo, estos datos no se consideran, ya que la información más reciente en COCHILCO es desde el año 2010, momento en que la gran minería comenzó un plan de producción sustentable con energías limpias. Además, con estos métodos es necesario una mayor cantidad de datos, y en este estudio solo se tomaron datos mensuales para la predicción. Debido a que la información diaria es difícil de controlar, y las empresas se muestran reacias a revelar esa información.

## 7. CONCLUSIONES

En este estudio, se construyeron modelos de redes neuronales recurrentes del tipo RNR, GRU y LSTM. El modelo propuesto puede ayudar a la gran minería del cobre a monitorear las características detrás de la producción de cobre y los insumos críticos utilizados para esto. Se considero un conjunto de datos en línea de observaciones de la producción de cobre y los insumos utilizados para su producción en Chile. El conjunto de datos incluye datos de producción de cobre, precio de la libra de cobre, consumo de agua, consumo de energía eléctrica y emisiones de GEI, desde el año 2000 al año 2020. Estos datos se utilizaron para pronosticar la producción de cobre en Chile.

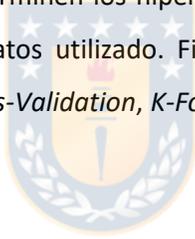
Fue posible realizar una predicción de la producción de cobre mediante técnicas de redes neuronales artificiales del tipo recurrentes, incluso con una baja cantidad de información. Sin embargo, el error y los niveles bajos de  $R^2$  obtenido, sugieren que es necesario una mayor cantidad de datos y nuevas configuraciones de estructuras de las redes. Por tanto, resultaría interesante mejorar el modelo ingresando nuevas variables como, por ejemplo, producción de ácido sulfúrico, producción de relaves y programación de mantenciones, con el fin de analizar y realizar proyecciones de producción de cobre más ajustadas.

Los resultados ilustran que el modelo de red neuronal RNR es un método simplista para pronosticar las tasas de producción de cobre en Chile; más específicamente, este modelo es incapaz de aprender las dependencias a largo plazo que causan el problema del descenso del gradiente, lo que impide que el algoritmo aprenda. El modelo GRU es adecuado para este conjunto de datos, ya que mostros un mejor ajuste que el modelo RNR, pero insuficientes porque, si bien presento bajos errores en las métricas, el  $R^2$  no demuestra un apropiado ajuste al conjunto de datos. Además, no hubo una mejora significativa en las métricas respecto al modelo RNR. El modelo LSTM da el mejor resultado en las métricas evaluadas, y las pruebas empíricas muestran que, el modelo puede lograr resultados factibles para proyectar la producción de cobre. Los hallazgos muestran que el modelo LSTM logra un alto ajuste a los datos, con valores  $R^2$  por sobre 0.7, y errores de hasta un 2% respecto a los datos reales.

En conclusión, los modelos implementados a partir de redes neuronales recurrentes, adaptados para proyectar la producción de cobre, presentan resultados factibles y apropiados con errores de hasta un 7% en todos los modelos experimentales. Resultan una herramienta útil para realizar proyecciones de producción minera y ayudar a obtener una visión profunda de la caracterización del yacimiento y las operaciones de producción.

Una alternativa de mejora para obtener un modelo más ajustado, al momento de proyectar la producción de cobre, sería incluir nuevas variables como: producción de ácido sulfúrico, toneladas de relaves producido, toneladas de escoria producida, programación de mantención de equipos y la inclusión de otros indicadores económicos, como el consumo mundial de cobre, PIB de los principales países compradores, entre otros.

Para estudios futuros se propone incorporar una visión económica de las proyecciones realizadas y disponer de una base de datos más específica y amplia de cada planta metalúrgica. Además, utilizar algoritmos que determinen los hiperparámetros y la estructura de la red, mejor ajustados para cada conjunto de datos utilizado. Finalmente incorporar distintos métodos de validación como: *Leave One Out Cross-Validation*, *K-Fold Cross-Validation* y *Repeated k-Fold-Cross-Validation*, entre otros.



## REFERENCIAS

- [1] USGS, (2020), "Material Commodity Summaries, US Geological Survey".
- [2] ALTALEY, (2020), "Hoja de Ruta 2.0 de la Minería Chilena Actualización y consensos para una mirada renovada".
- [3] Bata, M t H, Carriveau R, Ting, D S K (2020) "Short term water demand forecasting using nonlinear autoregressive artificial neural networks". Journal of Water Resources Planning and Management
- [4] S. Hochreiter, Y. Bengio, P. Frasconi, (2011). "Gradient flow in recurrent nets: The difficulty of learning long-term dependencies".
- [5] SERNAGEOMIN, (2019), "Anuario de la minería de Chile".
- [6] Banco Central, (2020), "Base de datos estadísticos".
- [7] F. Habashi, (1997), "Handbook of Extractive Metallurgy".
- [8] European Commission, (2015), "Report on critical raw materials for the EU".
- [9] COCHILCO, (2010), "Mercados e insumos estratégicos para la minería en Chile".
- [10] COCHILCO, (2015), "Factores clave para el desarrollo de la minería en Chile".
- [11] World Resources Institute, (2019-2020), "WRI Annual Report 2019-2020".
- [12] Organización de las Naciones Unidas, (2015), "Objetivos de desarrollo sostenible".
- [13] Ministerio de Obras Públicas, (2020), "Atlas de Calidad de Agua Chile".
- [14] COCHILCO, (2019), "Proyección del consumo de energía eléctrica en la minería del cobre 2019-2030.
- [15] Ministerio del Medio Ambiente, (2018), "Actualización de Chile sobre cambio climático".
- [16] I. Veličkovska, I. Mihajlović, B. Njagulović, "Prediction of the copper production in the framework of electrical energy consumption using artificial neural network", International May Conference on Strategic Management, (16) Issue (1) (2020) 411-423.
- [17] J. Baek, Y. Choi, "Deep Neural Network for Predicting Ore Production by Truck-Haulage Systems in Open-Pit Mines", Applied Sciences, (5), (2020) 1657.
- [18] N. Mamo, Y Dennis, "Artificial neural network based production forecasting for a hydrocarbon reservoir under water injection", Petroleum Exploration and Development, (47) Issue (2) (2020) 386-392.
- [19] I. Aizenberg, L. Sheremetov, L. Villa, J. Martinez, "Multilayer Neural Network with Multi-Valued Neurons in time series forecasting of oil production", Neurocomputing, (175) (2016) 980-989.
- [20] F. Almomani, "Prediction of biogas production from chemically treated co-digested agricultural waste using artificial neural network", Fuel, (280) (2020) 118573.
- [21] Y. Han, Z. Geng, Q. Zhu, "Energy optimization and prediction of complex petrochemical industries using an improved artificial neural network approach integrating data envelopment analysis", Energy Conversion and Management, (124) (2016) 73-83.

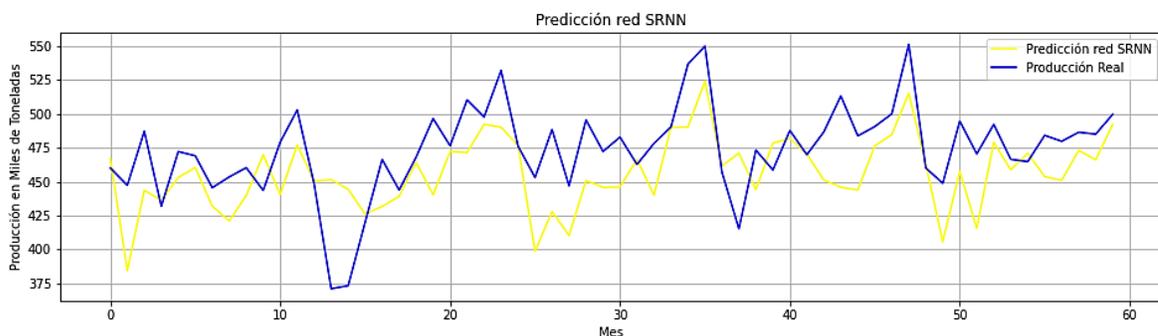
- [22] M. Bakay, Ü. Agbulut, "Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms", *Journal of Cleaner Production*, (285) (2021) 125324.
- [23] E. Uzlu, A. Akpınar, H. Öztürk, S. Nacar, M. Kankal, "Estimates of hydroelectric generation using neural networks with the artificial bee colony algorithm for Turkey", *Energy*, (69) (2014) 1-10.
- [24] R. Shabandar, A. Jaddoa, P. Liatsis, A. Hussain, "A deep gated recurrent neural network for petroleum production forecasting", *Machine Learning with Applications*, (3) (2021) 100013.
- [25] W. Liu, W. D. Liu, J. Gu, "Forecasting oil production using ensemble empirical model decomposition based Long Short-Term Memory neural network", *Journal of Petroleum Science and Engineering*, (189) (2020) 107013.
- [26] A. Sagheer, M. Kotb, "Time Series Forecasting of Petroleum Production using Deep LSTM Recurrent Networks", *Neurocomputing*, (3239) (2019) 203-213.
- [27] J. Puentes, C. Ribeiro, E. Ruelas, V. Figueroa, "Ethanol Fuel Demand Forecasting in Brazil Using a LSTM Recurrent Neural Network Approach", *IEEE Latin America Transactions*, (19) (2021)
- [28] G. Li, H. Wang, S. Zhang, J. Xin, H. Liu, "Recurrent Neural Networks Based Photovoltaic Power Forecasting Approach", *Energies*, (13) (2019) 2538.
- [29] Francois Chollet, (2018), "Deep Learning with Python".
- [30] M. Martínez, "Metodología basada en redes neuronales para interpretación de la resistividad del suelo en zonas urbanas", *Ingeniería Energética*, (1) (2014) 59-69.
- [31] F. Villada, N. Muñoz, E. García, "Redes Neuronales Artificiales aplicadas a la Predicción del Precio del Oro", *Información Tecnológica*, (27) (2016) 143-180.
- [32] C. Pan, J. Tan, "Day-ahead hourly forecasting of solar generation based on cluster Analysis and ensemble model", *IEEE Access*, (7) (2019) 112921–112930.
- [33] G. Cervone, L. Clemente, S. Alessandrini, L. Monache, "Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble", *Renew Energy*, (108) (2017) 274-286.
- [34] M. López, S. Valero, A. Rodriguez, I. Veiras, C. Senabre, "New online load forecasting system for the Spanish Transport System Operator", *Electric Power System Research*, (154) (2018) 401-412.
- [35] R. Monteiro, G. Guimaraes, F. Moura, M. Albertini, M. Albertini, "Estimating photovoltaic power generation: performance analysis of artificial neural networks, Support Vector Machine and Kalman filter", *Electric Power System Research*, (143) (2017) 643-656.
- [36] S. Vagropoulos, E. Kardakos, C. Simoglou, A. Bakirtzis, J. Catalao, "ANN-based scenario generation methodology for stochastic variables of electric power systems", *Electric Power System Research*, (134) (2016) 9-18.
- [37] Aurélien Géron, (2019), "Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow".
- [38] J. Bedi, D. Toshniwal, "Deep learning framework to forecast electricity demand", *Applied Energy*, (238) (2019) 1312-1326.

- [39] S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory", *Neural Computation*, (9) (1997) 1735-1780.
- [40] F. Karim, S. Majumdar, H. Darabi, S. Chen, "LSTM fully convolutional networks for time series classification", *IEEE Access*, (6) (2018) 1662–1669.
- [41] Y. Tian, L. Pan, "Predicting short-term traffic flow by long short-term memory recurrent neural network", *IEEE Access*, (2015).
- [42] S. Hochreiter, "Recurrent neural net learning and vanishing gradient", *International Journal Of Uncertainty*, (2) (1998) 107-116.
- [43] F. Weninger, J. Bergmann, B. Schuller, "Introducing Current: The Munich open-source Cuda Recurrent neural network toolkit", *Journal of Machine Learning Research*, (16) (2015) 547–551.
- [44] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, S. Valaee, "Recent advances in recurrent neural networks", (2017) 1–21.
- [45] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, "Learning phrase representations using RNN encoder–decoder for statistical machine translation. ArXiv Preprint", (2014).
- [46] J. De Gooijer, R. Hyndman, "25 years of time series forecasting", *International Journal of Forecasting*, (22) (2006) 443-473.
- [47] Q. Yang, X. Wu, "10 Challenging problems in data mining research", *International Journal of Information Technology and Decision Making*, (5) (2006) 597-604.
- [48] R. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, (50) (1982) 987-1007.
- [49] D. Erhan, C. Szegedy, A. Toshev, D. Anguelov, "Scalable object detection using deep neural networks", *IEEE Access*, (2014) 2147-2154.
- [50] M. Ebraheem, S. Thirumuruganathan, S. Joty, M. Ouzzani, N. Tang, "Distributed representations of tuples for entity resolution", *Proceedings of the VLDB Endowment*, (11) (2017) 1454-1467.

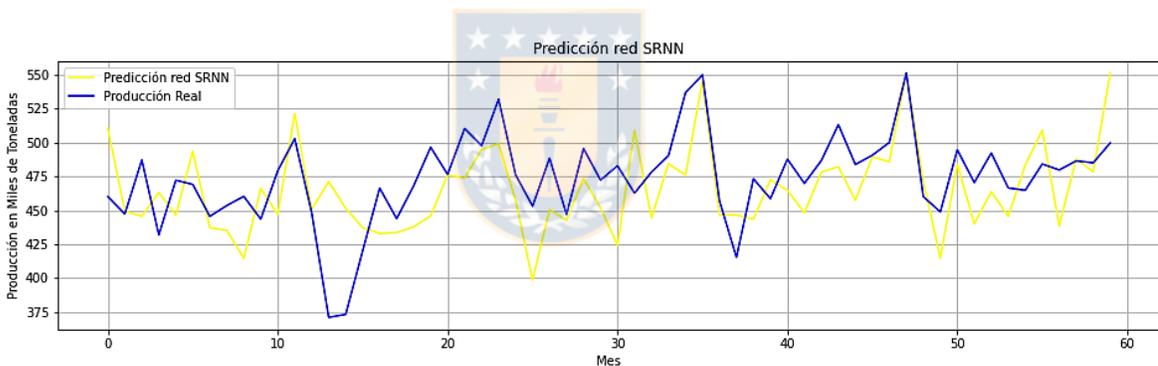
## A. ANEXOS

### A.1 Anexos modelo RNR

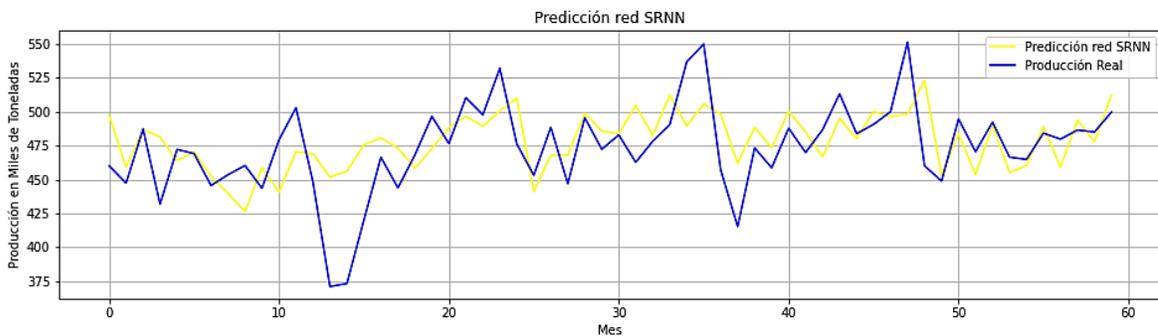
A continuación, se presentan las figuras de los resultados de predicción y entrenamiento del modelo RNR.



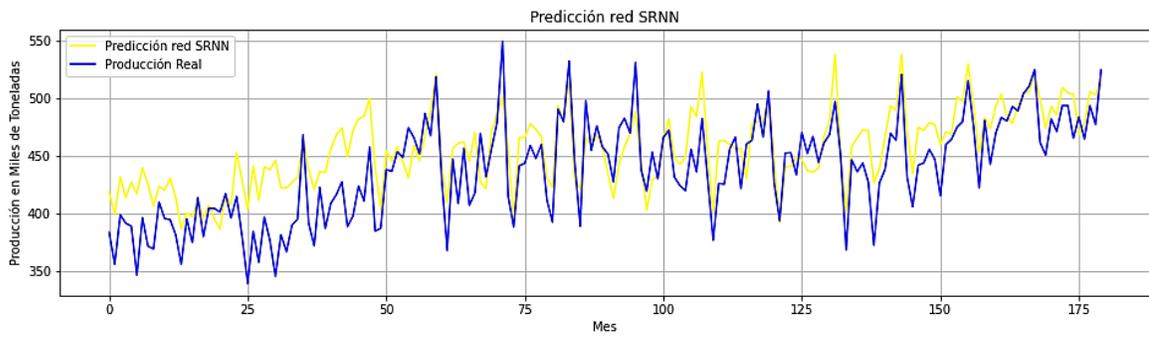
**Figura A.1** Grafica predicción modelo RNR (96) vs producción real.



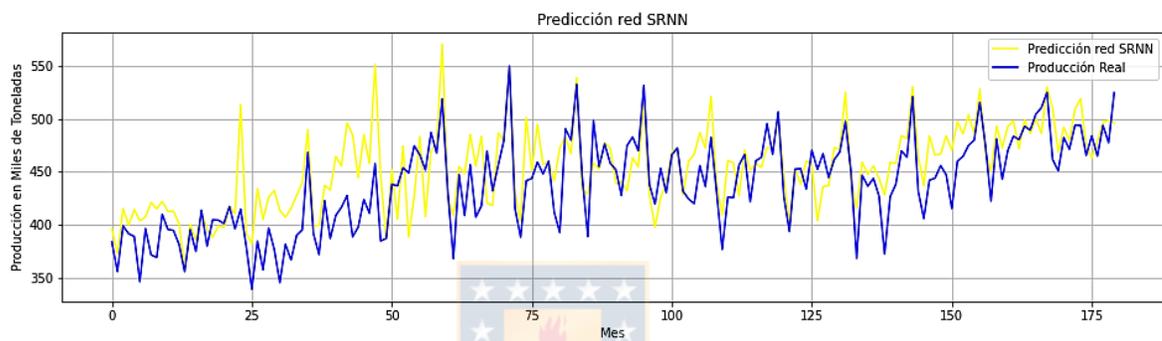
**Figura A.2.** Grafica predicción modelo RNR (64-96) vs producción real.



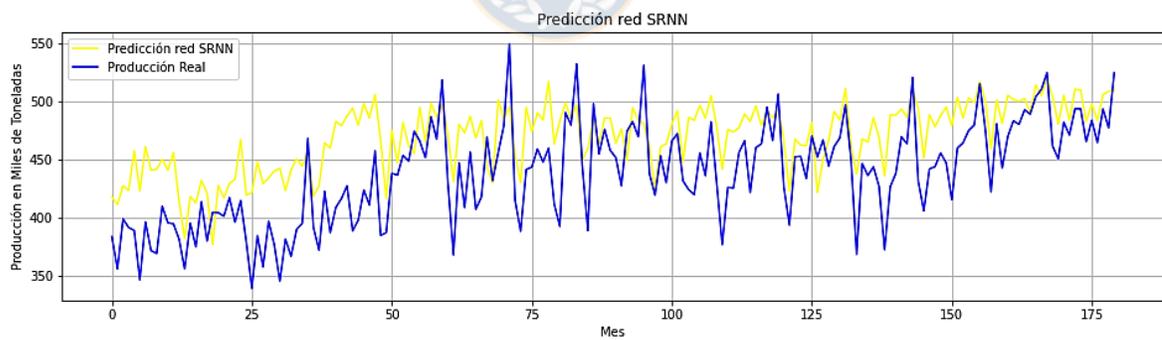
**Figura A.3.** Grafica predicción modelo RNR (64-96-96) vs producción real.



**Figura A.4.** Grafica entrenamiento modelo RNR (96) vs datos de entrenamiento.



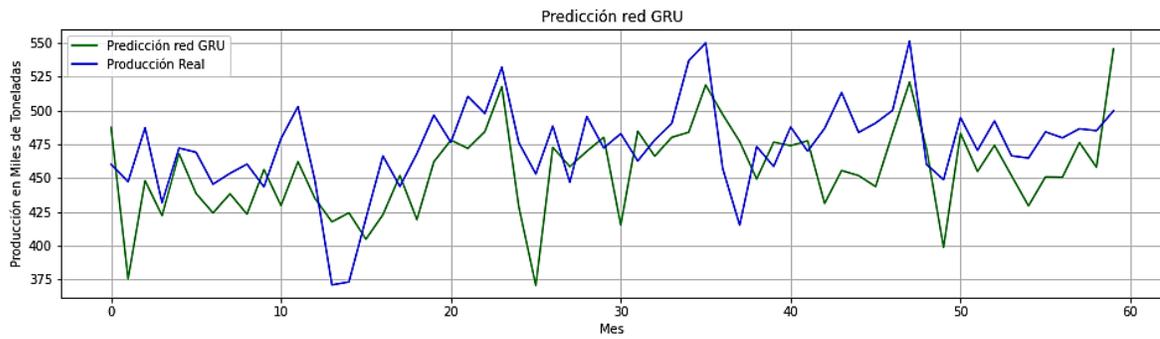
**Figura A.5.** Grafica entrenamiento modelo RNR (64-96) vs datos de entrenamiento.



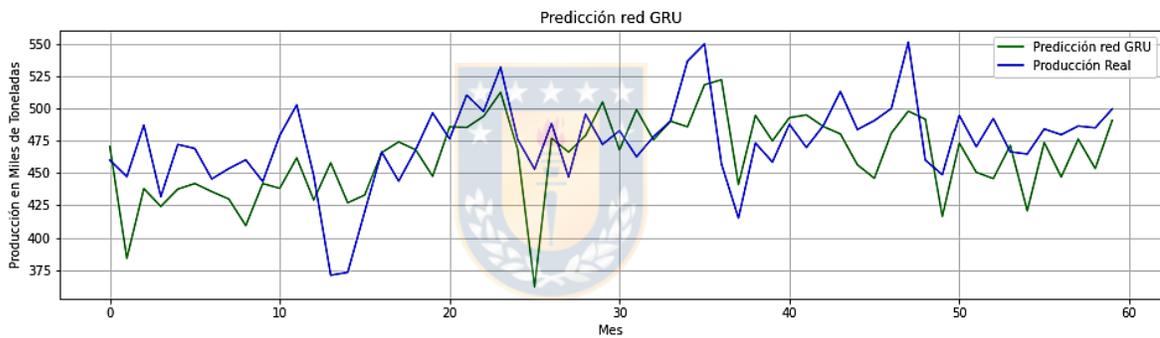
**Figura A.6.** Grafica entrenamiento modelo RNR (64-96-96) vs datos de entrenamiento.

## A.2 Anexos modelo GRU

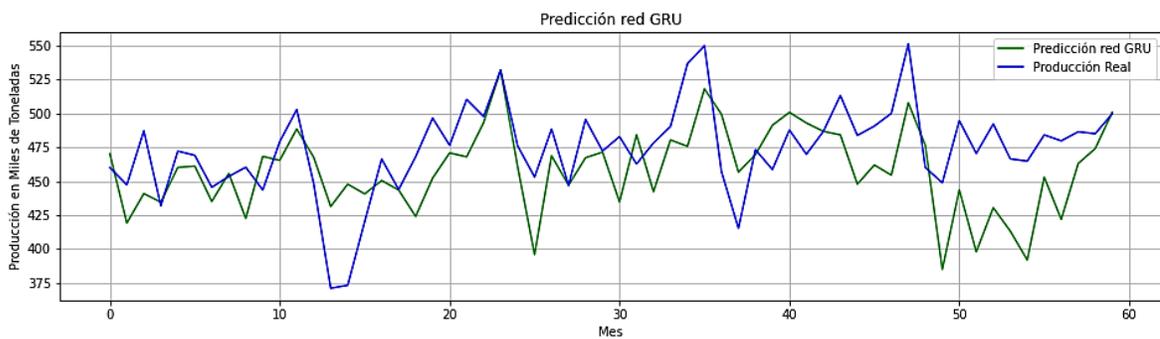
A continuación, se presentan las figuras de los resultados de predicción y entrenamiento del modelo GRU.



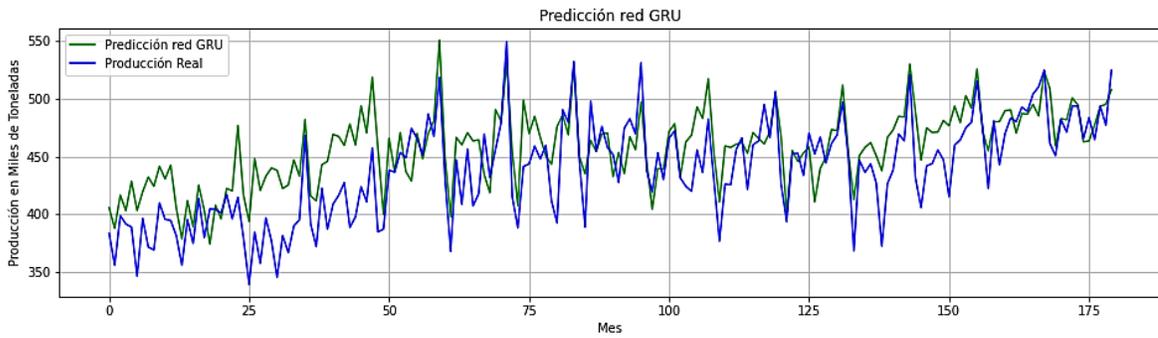
**Figura A.7.** Grafica predicción modelo GRU (96) vs producción real.



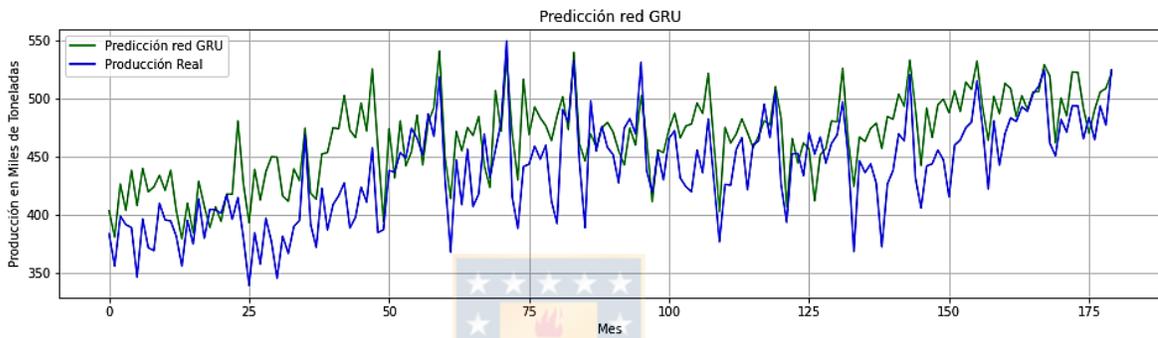
**Figura A.8.** Grafica predicción modelo GRU (64-64) vs producción real.



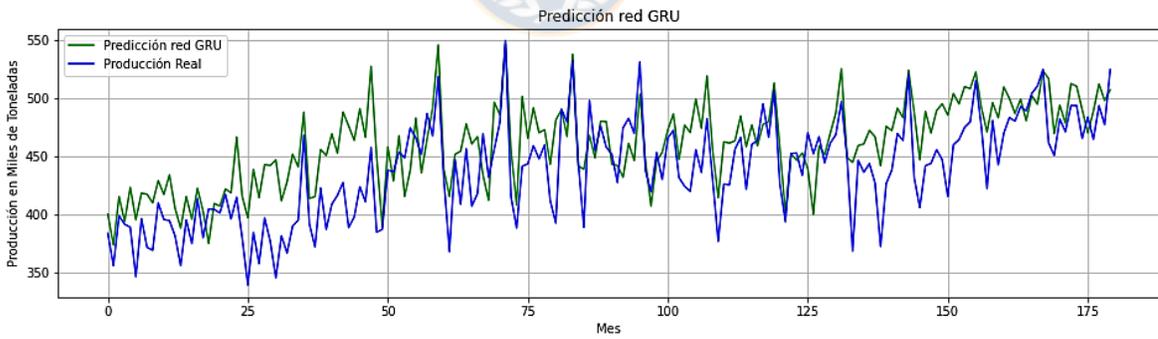
**Figura A.9.** Grafica predicción modelo GRU (64-96-96) vs producción real.



**Figura A.10.** Grafica entrenamiento modelo GRU (96) vs datos de entrenamiento.



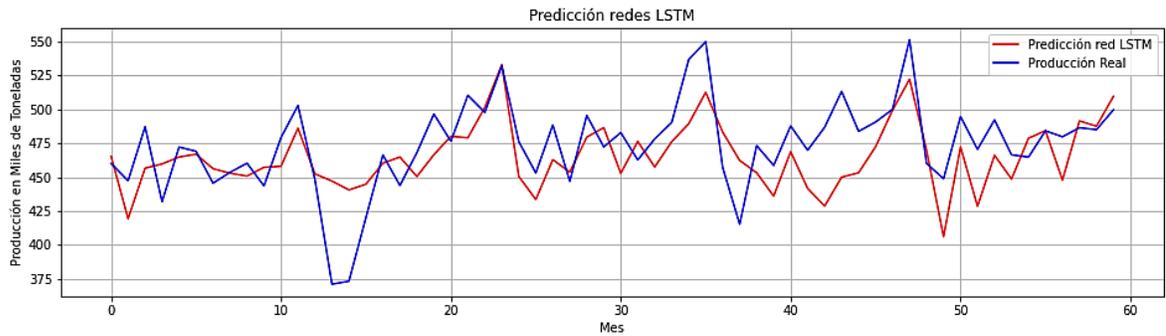
**Figura A.11.** Grafica entrenamiento modelo GRU (64-64) vs datos de entrenamiento.



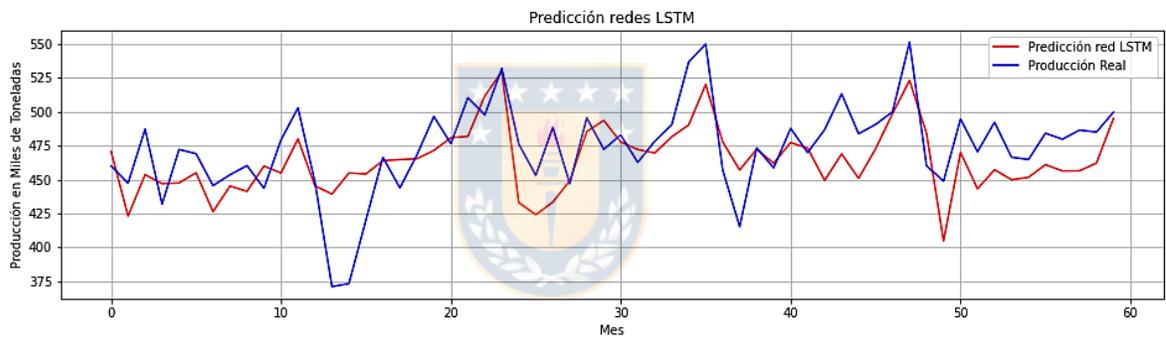
**Figura A.12.** Grafica entrenamiento modelo GRU (64-96-96) vs datos de entrenamiento.

### A.3 Anexos modelo LSTM

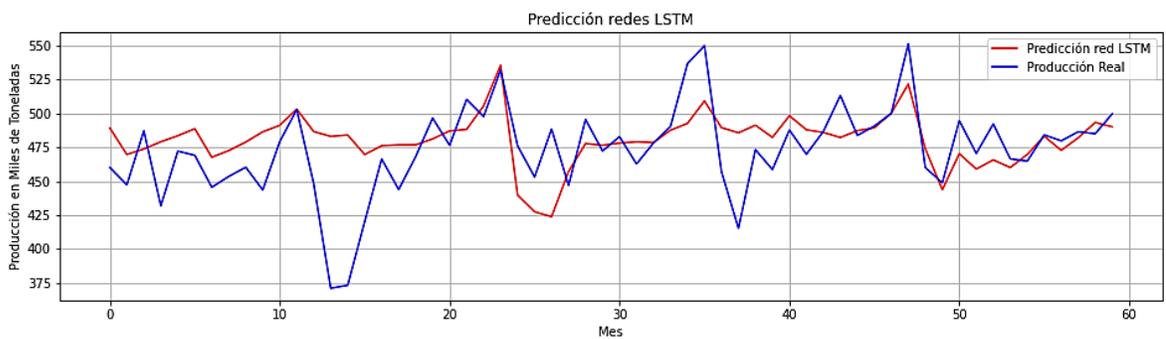
A continuación, se presentan las figuras de los resultados de predicción y entrenamiento del modelo LSTM.



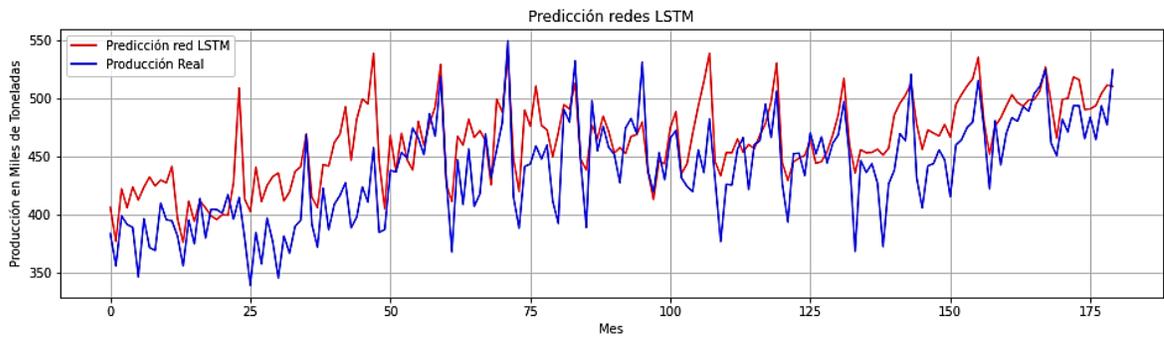
**Figura A.13.** Grafica predicción modelo LSTM (96) vs producción real.



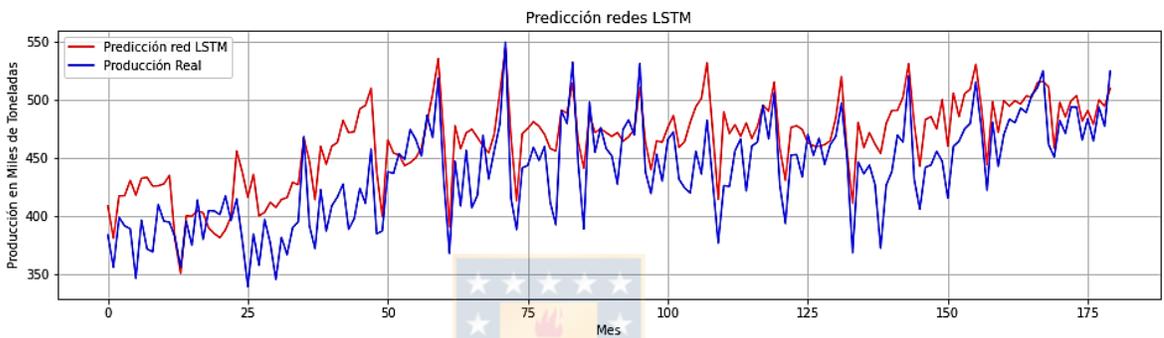
**Figura A.14.** Grafica predicción modelo LSTM (64-64) vs producción real.



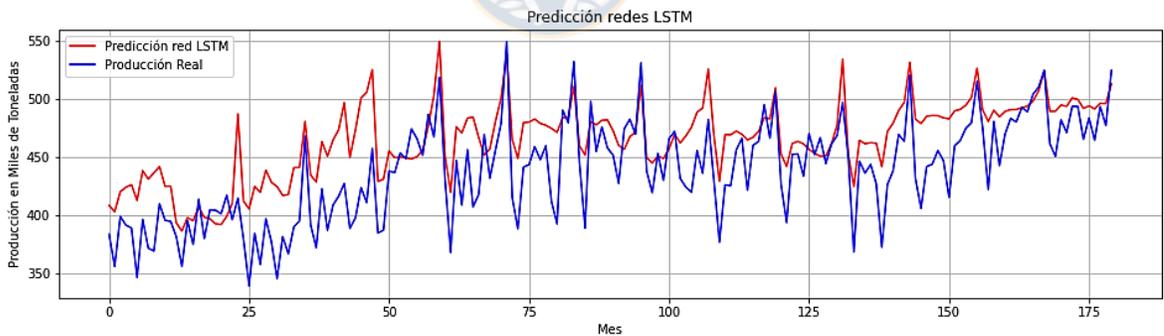
**Figura A.15.** Grafica predicción modelo LSTM (64-96-96) vs producción real.



**Figura A.16.** Grafica entrenamiento modelo LSTM (96) vs datos de entrenamiento.



**Figura A.17.** Grafica entrenamiento modelo LSTM (64-64) vs datos de entrenamiento.



**Figura A.18.** Grafica entrenamiento modelo LSTM (64-96-96) vs datos de entrenamiento.