



**UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL
CUSCO**

**ESCUELA DE POSGRADO
MAESTRÍA EN ESTADÍSTICA**

TESIS

**REDES NEURONALES EN LA PREDICCIÓN DE SERIES
TEMPORALES DE LAS DENUNCIAS POLICIALES DE LA
REGIÓN DEL CUSCO, 2015-2022**

**PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO EN
ESTADÍSTICA**

AUTOR:

Br. JESSICA HUANCACHOQUE COHAILA

ASESOR:

Dr. DIONICIO CARDENAS CANCHA

CÓDIGO ORCID:

0000-0003-0019-6659

CUSCO – PERÚ

2024

INFORME DE ORIGINALIDAD

(Aprobado por Resolución Nro.CU-303-2020-UNSAAC)

El que suscribe, asesor del trabajo de investigación/tesis titulado: REDES NEURONALES EN LA PREDICCIÓN DE SERIES TEMPORALES DE LAS DENUNCIAS POLICIALES DE LA REGIÓN DEL CUSCO, 2015-2022
presentado por: BI. JESSICA HUANDACHOQUE COHALLA

con Nro. de DNI: 41614781, para optar el título profesional/grado académico de MAESTRO EN ESTADÍSTICA

Informo que el trabajo de investigación ha sido sometido a revisión por 2 veces, mediante el Software Antiplagio, conforme al Art. 6° del *Reglamento para Uso de Sistema Antiplagio de la UNSAAC* y de la evaluación de originalidad se tiene un porcentaje de 10%

Evaluación y acciones del reporte de coincidencia para trabajos de investigación conducentes a grado académico o título profesional, tesis

Porcentaje	Evaluación y Acciones	Marque con una (X)
Del 1 al 10%	No se considera plagio.	X
Del 11 al 30 %	Devolver al usuario para las correcciones.	
Mayor a 31%	El responsable de la revisión del documento emite un informe al inmediato jerárquico, quien a su vez eleva el informe a la autoridad académica para que tome las acciones correspondientes. Sin perjuicio de las sanciones administrativas que correspondan de acuerdo a Ley.	

Por tanto, en mi condición de asesor, firmo el presente informe en señal de conformidad y **adjunto** la primera hoja del reporte del Sistema Antiplagio.

Cusco, 03 de JUNIO de 2024

Firma

Post firma DIONICIO CARDENAS CANOTA

Nro. de DNI 23840247

ORCID del Asesor 0000-0003-0019-6659

Se adjunta:

1. Reporte generado por el Sistema Antiplagio.
2. Enlace del Reporte Generado por el Sistema Antiplagio: oid:27259:359008821

NOMBRE DEL TRABAJO

REDES NEURONALES EN LA PREDICCIÓN DE SERIES TEMPORALES DE LAS DENUNCIAS POLICIALES DE LA REGIÓN DEL

AUTOR

JESSICA HUANCACHOQUE

RECUENTO DE PALABRAS

22598 Words

RECUENTO DE CARACTERES

126108 Characters

RECUENTO DE PÁGINAS

103 Pages

TAMAÑO DEL ARCHIVO

2.2MB

FECHA DE ENTREGA

Jun 3, 2024 12:04 PM GMT-5

FECHA DEL INFORME

Jun 3, 2024 12:09 PM GMT-5**● 10% de similitud general**

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos.

- 9% Base de datos de Internet
- Base de datos de Crossref
- 6% Base de datos de trabajos entregados
- 1% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

● Excluir del Reporte de Similitud

- Material bibliográfico
- Material citado
- Bloques de texto excluidos manualmente
- Material citado
- Coincidencia baja (menos de 20 palabras)

DEDICATORIA

Dedico esta tesis a mi padre celestial mi Dios, por ser mi guía en mi vida, a mi amada madre Rosaura por siempre orar por mí y sus bendiciones me protegen cada día, gracias, madre por darme la vida, a mi querido Padre Jacinto por siempre ser el ejemplo a seguir, luchador y perseverante, a mis hermanos por brindarme su apoyo en todo momento, a mis queridos sobrinos por enseñarme a ser una gran persona.

Dedico esta tesis a mi amado Percy por siempre brindarme su apoyo y amor incondicional, sus consejos siempre los tengo presente, a mis amados hijos Leandro y José por siempre contar con ellos, por ser mi fuente de inspiración y motivación para poderme superar cada día más.

AGRADECIMIENTOS

En primer lugar, agradecer a mi Creador, quien es mi fortaleza y que siempre estuvo, está y estará a mi lado a lo largo de mi vida, gracias DIOS, de gracia nació doy gracias, en segundo lugar, agradecer a mis queridos padres por las enseñanzas y valores que recibí, por siempre brindarme su apoyo en todo momento de mi vida, la gran familia que tengo; a mi amado esposo por siempre confiar y guiarme en las decisiones importantes de seguir siempre capacitándome y actualizándome, a mis queridos hijos Leandro y Jose por ser la inspiración de seguir adelante, gracias a ellos conocí el verdadero amor puro y sincero.

Agradecer a mi asesor Dionicio Cárdenas y a todas las personas que hicieron posibles el logro de esta tesis por sus consejos, guías y siempre alentarme a seguir adelante, a todos mis amigos que, con una palabra de aliento hicieron que en mí surja la necesidad de seguir luchando.

Agradezco a mi amada Universidad San Antonio Abad del Cusco por la enseñanza, la buena experiencia, por permitirme convertirme en ser una gran profesional que se apasiona con su carrera, a cada docente que fue parte de este proceso de mi formación, el cual deja como producto terminado y prueba viviente en la historia, esta tesis que trascenderá dentro de los conocimientos y desarrollo de las demás generaciones que están por llegar.

Finalmente agradezco a quien lee este apartado, por permitir que mis experiencias, investigación y conocimiento llene su repertorio de información.

INDICE DE CONTENIDO

DEDICATORIA	ii
AGRADECIMIENTOS	iii
INDICE DE CONTENIDO	iv
INDICE DE TABLAS	vii
INDICE DE FIGURAS	vii
RESUMEN	ix
RESUMO	x
CAPÍTULO I	1
PLANTEAMIENTO DEL PROBLEMA	1
1.1. DESCRIPCIÓN DEL PROBLEMA DE INVESTIGACIÓN	1
1.2. FORMULACIÓN DEL PROBLEMA	3
1.2.1. PROBLEMA GENERAL	3
1.2.2. PROBLEMA ESPECIFICOS	3
1.3. OBJETIVOS DE LA INVESTIGACION	3
1.3.1. OBJETIVO GENERAL	3
1.3.2. OBJETIVOS ESPECIFICOS	4
1.4. JUSTIFICACIÓN DE LA INVESTIGACIÓN	4
CAPÍTULO II	6
MARCO TEORICO CONCEPTUAL	6
2.1. ANTECEDENTES	6
2.1.1. ANTECEDENTES INTERNACIONALES	6
2.1.2. ANTECEDENTES NACIONALES	8
2.2. BASES TEÓRICAS	9
2.2.1. DENUNCIA	9
2.2.2. DENUNCIANTE	13
2.2.3. DENUNCIADO	14
2.3. REDES NEURONALES ARTIFICIALES	14
2.3.1. REDES NEURONALES BIOLÓGICAS	14
2.3.2. REDES NEURONALES RECURRENTE Y LSTM	33
2.4. SERIES TEMPORALES	39
2.4.1. COMPONENTES DE UNA SERIE TEMPORAL	39
2.4.2. CLASIFICACIÓN DESCRIPTIVA DE LAS SERIES TEMPORALES	40
2.4.3. PROCESO ESTOCÁSTICO ESTACIONARIO	40
2.4.4. RUIDO BLANCO	41
2.4.5. FUNCIÓN DE AUTOCORRELACIÓN	41
2.4.6. MODELOS DETERMINÍSTICOS	41
2.5. METODOLOGÍA BOX JENKINS	42

2.5.1. MODELOS DE PROMEDIO MÓVIL AUTORREGRESIVO INTEGRADO: ARIMA (P,D,Q).....	43
2.5.2. MODELOS ARIMA ESTACIONALES	45
2.5.3. PRUEBAS PARA ANALIZAR LA TENDENCIA.....	46
2.5.4. ANÁLISIS DE LA ESTACIONALIDAD MEDIANTE GRÁFICOS.....	50
2.5.5. EVALUACIÓN DE LA BONDAD DEL MODELO PREDICTIVO	51
2.6. HIPOTESIS	52
2.6.1. HIPOTESIS GENERAL.....	52
2.6.2. HIPOTESIS ESPECIFICAS	52
2.7. IDENTIFICACION DE VARIABLES E INDICADORES	53
2.8. OPERACIONALIZACIÓN DE VARIABLES.....	53
CAPÍTULO III.....	55
METODOLOGÍA	55
3.1. ÁMBITO DE ESTUDIO: LOCALIZACIÓN POLÍTICA Y GEOGRÁFICA	55
3.2. TIPO Y NIVEL DE INVESTIGACIÓN.....	56
3.2.1. ENFOQUE DE INVESTIGACIÓN	56
3.2.2. TIPO DE INVESTIGACIÓN.....	56
3.2.3. ALCANCE O NIVEL DE INVESTIGACIÓN	56
3.2.4. DISEÑO DE INVESTIGACIÓN	56
3.3. UNIDAD DE ANALISIS.....	56
3.4. POBLACION DE ESTUDIO.....	56
3.5. TAMAÑO DE MUESTRA	56
3.6. TECNICAS DE RECOLECCION DE INFORMACIÓN.....	56
3.7. TECNICAS DE ANALISIS E INTERPRETACIÓN DE LA INFORMACION.....	57
CAPÍTULO IV	58
RESULTADOS Y DISCUSIÓN.....	58
4.1. ANÁLISIS DE LA COMPONENTE DE TENDENCIA EN LAS DENUNCIAS POLICIALES EN LA REGIÓN DEL CUSCO, 2015-2022.....	58
4.1.1. ANÁLISIS DE LA TENDENCIA: MANN KENDALL	59
4.1.2. ANÁLISIS DE LA TENDENCIA: COX STUART	60
4.1.3. ANÁLISIS DE LA TENDENCIA: CORRELACIÓN DE SPEARMAN	60
4.1.4. ANÁLISIS DE LA TENDENCIA: WALD WOLFOWITZ.....	60
4.2. ANÁLISIS DE LA COMPONENTE ESTACIONALIDAD EN LAS DENUNCIAS POLICIALES EN LA REGIÓN DEL CUSCO, 2015-2022.....	61
4.2.1. ANÁLISIS DE LA ESTACIONALIDAD: MONTHPLOT	61
4.2.2. ANÁLISIS de LA ESTACIONALIDAD: BOXPLOT.....	61
4.2.3. ANÁLISIS de la ESTACIONALIDAD: USO DEL CORRELOGRAMA	62
4.2.4. ANÁLISIS DE LA ESTACIONALIDAD: PRUEBA DE KRUSKAL WALLIS	63
4.3. METODOLOGÍA BOX JENQUINS – ARIMA(p,d,q)	63
4.3.1. FASE DE IDENTIFICACIÓN	63

4.3.2. FASE DE ESTIMACIÓN DE LOS MODELOS	66
4.3.3. FASE DE DIAGNÓSTICO Y VALIDACIÓN.....	69
4.3.4. FASE DE PREDICCIÓN ARIMA (1,1,0)	72
4.4. APLICACIÓN DE REDES NEURONALES EN SERIES TEMPORALES DE LAS DENUNCIAS POLICIALES EN LA PROVINCIA DEL CUSCO, 2015-2022.	73
4.4.1. FASE DE AJUSTE DEL MODELO	73
4.4.2. PRECISIÓN Y VALIDACIÓN DEL MODELO.....	73
4.4.3. FASE DE PREDICCIÓN NNAR (1,1,2) [12].....	76
4.5. DISCUSIÓN DE RESULTADOS	77
CAPÍTULO V	79
CONCLUSIONES Y RECOMENDACIONES	79
CONCLUSIONES	79
RECOMENDACIONES	81
REFERENCIAS BIBLIOGRÁFICAS.....	82
ANEXOS.....	85
Anexo 1. MATRIZ DE CONSISTENCIA	86
Anexo 2. DATOS PROPORCIONADOS PNP	88
Anexo 3. FICHA DE RECOLECCIÓN DE DATOS	90
Anexo 4. CÓDIGO DE R	91
Anexo 5. DATOS.....	95
Anexo 6. SOLICITUDES DE INFORMACIÓN	96

INDICE DE TABLAS

Tabla 1. Funciones de Transferencia.....	18
Tabla 2. Operacionalización de variable	53
Tabla 3. Comparación de modelos ARIMA.....	68
Tabla 4. Comparación de Precisión y validación de los modelos ARIMA Y Red Neuronal.....	74
Tabla 5. Matriz de consistencia	86

INDICE DE FIGURAS

Figura 1. Estructura del proceso de una denuncia	13
Figura 2. Modelo simplificado de una neurona biológica.....	15
Figura 3. Modelo de neurona estándar.....	16
Figura 4. Esquema de una Neurona Simple.....	17
Figura 5. Representación de un Clasificador.....	21
Figura 6. Representación de un perceptrón	23
Figura 7. Ilustración del método de gradiente descendiente	24
Figura 8. Regiones de Decisión.....	26
Figura 9. Representación de una Red Neuronal Multicapa	26
Figura 10. Ilustración de Superficie de Error	32
Figura 11. Ejemplo de una red neuronal recurrente.....	33
Figura 12. Ejemplo de una celda LSTM	36
Figura 13. Rendimiento de modelos	38
Figura 14. Mapa del departamento del Cusco	55
Figura 15. Gráfico de la serie de tiempo de las denuncias	58
Figura 16. Monthplot.....	61
Figura 17. Gráfico de Cajas por meses.....	61
Figura 18. Función de autocorrelación	62
Figura 19. Autocorrelación.....	65
Figura 20. Autocorrelación parcial.....	65
Figura 21. Residuales ARIMA (1,1,0)	69
Figura 22. Residuales ARIMA(1,1,1)	69

Figura 23. Autocorrelación de los residuales- ARIMA (1,1,0).....	70
Figura 24. Autocorrelación de los residuales-ARIMA(1,1,1).....	71
Figura 25. Predicción con el modelo ARIMA(1,1,0)	72
Figura 26. Residuales - NNAR(1,1,2)[12].....	75
Figura 27. Autocorrelación de los residuales- NNAR(1,1,2)[12].....	75
Figura 28. Proyecciones con el modelo NNAR(1,1,2)[12].....	76

RESUMEN

La problemática de la inseguridad ciudadana se manifiesta como un fenómeno y una cuestión de índole social en estos tiempos, afectando a la sociedad en su desarrollo económico, aspectos culturales y patrones comportamentales; ahí la necesidad de observar el comportamiento de las denuncias, para los próximos meses del 2023 mediante las redes neuronales en series temporales de las denuncias policiales de la región del Cusco. El **método** de estudio es aplicado, descriptivo, con diseño de investigación no experimental, transversal. Se analizaron los modelos de redes neuronales recurrentes para series temporales LSTM (Long Short Term Memory), además de modelos ARIMA mediante la metodología de box y Jenkins; los **resultados** a los que se llegaron fueron: Existe tendencia creciente, así mismo se comprobó que no está presente la componente de estacionalidad, y mediante indicadores de precisión como el AIC, RMSE, MAE, MPE, MAPE de los modelo ARIMA (1,1,1), el modelo ARIMA (1,1,0), y NNAR(1,1,2)[12], el modelo de redes neuronales recurrentes LSTM NNAR(1,1,2)[12] es el que presenta mejores parámetros de ajuste; la **conclusión**: La red neuronal recurrentes LSTM NNAR(1,1,2)[12] presentan un mejor ajuste a la serie de tiempo de las denuncias policiales en la región del Cusco la cual fue producto de 20 redes, presenta dos capas ocultas ($k=2$) el error medio ($ME=0.0502$); así mismo la raíz del cuadrado medio del error (RMSE), el error medio absoluto ($MAE =105.375$), MAPE (Error Porcentual Absoluto Medio) también presento mejor ajuste con las redes neuronales que con los otros modelos comparados. Para el año 2023, considerando la red neuronal para series de tiempo NNAR (1,1,2) [12], para los 3 últimos meses del año 2023 se estima que el número de denuncias por mes sean de 1299.55, 1302.46 y 1302.403.

Palabras clave: Redes neuronales recurrentes, series de tiempo, denuncias.

RESUMO

O surgimento da insegurança do cidadão se manifesta como um fenômeno e uma questão social nos tempos atuais, afetando a sociedade em seu desenvolvimento econômico, aspectos culturais e padrões de comportamento; daí a necessidade de observar o comportamento das queixas para os próximos meses de 2023 por meio de redes neurais em séries temporais de queixas policiais na região de Cusco. O estudo é aplicado, descritivo, e o desenho da pesquisa é não experimental, longitudinal. E as técnicas estatísticas a serem analisadas são os modelos de redes neurais recorrentes para séries temporais LSTM (Long Short Term Memory), além de modelos ARIMA usando a metodologia box e Jenkins; os resultados alcançados foram: Existe tendência de aumento, da mesma forma verificou-se que o componente de sazonalidade não está presente, e por meio de indicadores de precisão como o AIC, RMSE, MAE, MPE, MAPE dos modelos ARIMA (1,1,1), o modelo ARIMA (1,1,0), e NNAR(1,1,2)[12], dá por justo vencedor ao modelo de redes neurais recorrentes LSTM NNAR(1,1,2)[12] que apresenta melhores parâmetros de ajuste; a conclusão geral da pesquisa foi que a rede neural recorrente LSTM NNAR(1,1,2)[12] apresenta um melhor ajuste à série temporal de boletins de ocorrência na região de Cusco, que foi o produto de 20 redes, apresenta duas camadas ocultas ($k=2$), o erro médio ($ME=0.0502$); da mesma forma, a raiz do erro quadrático médio (RMSE), o erro absoluto médio ($MAE=105,375$), o MAPE (erro percentual absoluto médio) também apresentaram melhor ajuste com as redes neurais do que com os outros modelos comparados. Para o ano de 2023, considerando a rede neural para séries temporais NNAR (1,1,2) [12], para os últimos 3 meses do ano de 2023, são esperadas estimativas de 1299,55, 1302,46 e 1302,403 reclamações, respectivamente.

Palavras-chave: Redes neurais recorrentes, séries temporais, denúncias.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1. DESCRIPCIÓN DEL PROBLEMA DE INVESTIGACIÓN

La emergencia de la inseguridad ciudadana se manifiesta como un fenómeno y una cuestión de índole social en la contemporaneidad, afectando a sociedades con una amplia diversidad en su desarrollo económico, marcadas por una pluralidad de aspectos culturales y gobernadas por diversos regímenes políticos. En este contexto, no es viable la formulación de categorizaciones simplistas para la identificación de los factores vinculados a su crecimiento y las modalidades en que se manifiesta (Nuñez, 2015).

En la actualidad, La percepción de seguridad pública es baja o nula para la mayoría de los peruanos, ya que el crimen se ha convertido en una preocupación importante tanto para los ciudadanos como para las autoridades (INEI, 2018). Para comprobarlo, basta con mirar las noticias que suceden en la calle todos los días y darse cuenta de que esta percepción es cierta (Cano, 2023).

Indudablemente, la crisis en el ámbito de la seguridad y la consiguiente imperativa exigencia de reformas han ocasionado que la población emprenda acciones autónomas en busca de justicia y protección, valiéndose de estrategias como las campañas en las redes sociales y enfrentando de manera directa la amenaza mortal (Nuñez, 2015).

Conforme a la información encontrada en el Instituto Nacional de Estadística e Informática (INEI), los indicadores de actividad criminal están retomando los niveles registrados antes de la pandemia. Este fenómeno se atribuye, entre otros factores, al levantamiento de las restricciones de movilidad, la escasez de oportunidades laborales y la implementación de políticas efectivas. Las estadísticas relacionadas con la criminalidad ponen de manifiesto que el país aún enfrenta un significativo desafío en términos de seguridad ciudadana, que incide directamente en el derecho fundamental de cada individuo a la vida, la libertad y la integridad. La criminalidad se manifiesta a través de una variedad de modalidades delictivas, todas ellas reflejadas en las denuncias presentadas ante las autoridades policiales (INEI, 2021).

Según el más reciente informe sobre estadísticas de seguridad ciudadana para el periodo septiembre 2022-febrero 2023 (INEI, 2023b), la proporción de la población urbana de 15 años o más que fue víctima de algún hecho delictivo fue del 23.2%. Esta cifra representa un aumento de 0.7 puntos porcentuales en comparación con el periodo julio 2022-diciembre 2022 y de 3.3 puntos porcentuales en relación con el mismo periodo de 2021. Es relevante destacar que este resultado está 6.4 puntos porcentuales por encima del índice más bajo registrado durante el periodo de la pandemia, que fue durante marzo 2021-agosto 2021 (INEI, 2023a).

Es destacable señalar que, en la población urbana a nivel nacional, los individuos de 15 a 29 años constituyen el grupo de edad más impactado por la actividad delictiva, ya que el 26.9% de este grupo informó haber sido víctima de delitos durante el año 2022. Luego se sitúan los adultos entre 30 y 49 años, con una incidencia del 25.4%, mientras que en los adultos entre 50 y 59 años fue del 22.1%. Los adultos mayores, de 60 años a más, fueron los menos afectados, con el 11.9% (INEI, 2023b).

Es de relevancia destacar que, en el contexto de la población urbana a nivel nacional, el segmento poblacional más susceptible a los efectos de la delincuencia corresponde a los jóvenes, específicamente aquellos con edades comprendidas entre los 15 y los 29 años, con un porcentaje del 26.9% de estos que informaron haber experimentado situaciones victimizantes durante el año 2022. Subsecuentemente, se encuentra el grupo de adultos de 30 a 49 años, con una incidencia del 25.4%, seguido de los adultos que oscilan entre los 50 y los 59 años, quienes presentaron una tasa del 22.1%. En contraste, los adultos mayores, es decir, aquellos de 60 años en adelante, exhibieron la menor incidencia, con un 11.9% (INEI, 2023b).

Según el Instituto Nacional de Estadística e Informática (INEI), la abrumadora mayoría de las denuncias registradas corresponden a infracciones contra el patrimonio, que engloban una variedad de delitos como robos y estafas, entre otros. Estos delitos representaron el 66.4% del total de denuncias en los últimos meses del año 2022. Las regiones más afectadas por este tipo de incidentes fueron Ucayali, con un notable 76.7% de los delitos concentrados en esta categoría, seguido de Lima Metropolitana y Piura, ambos con un 74.1% (INEI, 2023b).

Estos patrones observados en las denuncias a nivel nacional indican que la formulación de los planes de prevención del delito se está llevando a cabo de manera inadecuada, lo que contribuye a la situación actual que se experimenta (Reyna & Toche, 1999). Se nota un aumento en la incidencia de delitos a lo largo de los años,

también es importante resaltar el paralelo avance de la tecnología, particularmente en el campo de la inteligencia artificial. Gracias a este progreso tecnológico, las fuerzas de seguridad en otros países han logrado mejorar su capacidad de respuesta frente a la delincuencia. En el contexto de nuestro país, Perú, se alberga la esperanza de que se pueda incorporar la inteligencia artificial en este caso en particular algoritmos de redes neuronales como herramienta para predecir la cantidad de delitos. Este enfoque busca estimular a las autoridades a desarrollar políticas que aborden de manera efectiva esta problemática que impacta profundamente en la sociedad.

1.2. FORMULACIÓN DEL PROBLEMA

1.2.1. PROBLEMA GENERAL

¿Cuáles son los modelos de redes neuronales que mejor predicen series temporales de las denuncias policiales de la región del Cusco, 2015-2022?

1.2.2. PROBLEMA ESPECIFICOS

- a) ¿Cuál es la evolución temporal de las denuncias policiales en la región del Cusco durante el período comprendido entre 2015 y 2022, en su componente de la tendencia?
- b) ¿Cuál es el patrón estacional observado en las denuncias policiales en la región del Cusco durante el periodo de 2015 a 2022?
- c) ¿Cuál es el modelo ARIMA que mejor ajusta a la serie de datos denuncias policiales en la provincia del Cusco, 2015-2022, para la predicción de los 3 últimos meses del año 2023?
- d) ¿Cuál es el modelo de redes neuronales en series temporales que mejor predice a la serie de datos de las denuncias policiales en la provincia del Cusco, de los 3 últimos meses del año 2023?

1.3. OBJETIVOS DE LA INVESTIGACION

1.3.1. OBJETIVO GENERAL

Determinar los modelos de redes neuronales en series temporales que mejor predicen las denuncias policiales en la Región del Cusco, 2015-2022

1.3.2. OBJETIVOS ESPECIFICOS

- a) Analizar la evolución temporal de las denuncias policiales en la región del Cusco durante el período 2015-2022, con respecto de la componente tendencia.
- b) Determinar el patrón estacional presente en las denuncias policiales en la región del Cusco durante el periodo comprendido entre 2015 y 2022.
- c) Determinar el modelo ARIMA que mejor se ajusta a la serie temporal de datos denuncias policiales en la provincia del Cusco durante el periodo de 2015 a 2022
- d) Determinar el modelo de redes neuronales en series temporales que mejor predice a la serie de datos de las denuncias policiales en la provincia del Cusco durante el periodo de 2015 a 2022

1.4. JUSTIFICACIÓN DE LA INVESTIGACIÓN

El análisis de las denuncias policiales en la Región del Cusco es una cuestión de vital importancia que merece una atención cuidadosa y exhaustiva. Esta justificación se basa en varias razones fundamentales que resaltan la necesidad de abordar esta problemática:

La investigación de las denuncias policiales reviste una importancia fundamental en el análisis de las dinámicas sociales y comunitarias en la Región del Cusco. Dichas denuncias no solo constituyen un reflejo de las inquietudes de la ciudadanía, sino que también pueden poner de manifiesto problemáticas subyacentes que inciden en la calidad de vida y la seguridad de la población.

Examinar las denuncias policiales puede conferir datos de relevancia sustancial acerca de los tipos de infracciones más frecuentes en la región. Esta información reviste una importancia crucial para que las autoridades y las fuerzas del orden puedan adoptar medidas proactivas destinadas a mejorar la seguridad ciudadana y prevenir la comisión de delitos.

La evaluación de las denuncias policiales posibilita la identificación de tendencias delictivas, patrones de criminalidad y períodos de mayor incidencia delictiva. Estos datos revisten una importancia crucial en la formulación de políticas públicas eficaces y estrategias de aplicación de la ley. Los resultados de este análisis pueden constituir un fundamento sustantivo para investigaciones académicas ulteriores y proyectos

vinculados a la seguridad ciudadana y la ejecución de la ley en la región; así mismo con esta investigación fomentamos la transparencia y la rendición de cuentas en las instituciones policiales y judiciales, además de facilitar la evaluación de los procesos de gestión y resolución de denuncias, aspecto que, a su vez, incide en el fortalecimiento de la confianza comunitaria en las instituciones responsables de la aplicación de la ley.

El incremento de la seguridad y la disminución de la delincuencia ejercen un impacto inmediato en la calidad de vida de los habitantes de la Región del Cusco. La reducción de la actividad criminal puede propiciar un entorno más seguro y propicio para el desarrollo económico y social.

CAPÍTULO II

MARCO TEORICO CONCEPTUAL

2.1. ANTECEDENTES

2.1.1. ANTECEDENTES INTERNACIONALES

En su tesis de investigación titulada "**Aplicación de Redes Neuronales para el Pronóstico de Series de Tiempo: Exploración de las Series Temporales Avanzadas**" (Fernandez, 2021), el autor se propuso principalmente demostrar la utilidad de las redes neuronales en el contexto de análisis de series de tiempo. Se destacó que, a pesar de las ventajas de las redes neuronales en ciertos escenarios, los modelos tradicionales siguen siendo viables y menos complejos en diferentes contextos, ya que proporcionan una mayor capacidad explicativa en términos de mecanismo. Esto se debe a que muchas redes neuronales se perciben como cajas negras, lo que dificulta la interpretación de sus resultados.

Los investigadores llevaron a cabo una evaluación exhaustiva del rendimiento predictivo del aprendizaje profundo en diversas situaciones relacionadas con series temporales. Concluyeron que, aunque el aprendizaje profundo puede mejorar la precisión de las predicciones cuando se dispone de grandes conjuntos de datos, no es necesariamente la solución óptima en todos los contextos. En situaciones más simples o en problemas que requieren soluciones analíticas específicas debido a su naturaleza, los modelos tradicionales pueden ser más eficaces (Fernandez, 2021). Este estudio aporta una comprensión más profunda del uso de las redes neuronales en el pronóstico de series de tiempo, subrayando la importancia de elegir la metodología adecuada según la complejidad y los requisitos específicos de cada problema (Fernandez, 2021).

Agustín, en el (2017) realizó una tesis de investigación la cual intitula "**Redes Neuronales Recurrentes: Una aplicación para los mercados bursátiles**" en la búsqueda de anticipar las tendencias de una destacada empresa del índice S&P 500, EBAY, durante un horizonte de 15 días, se emplearon avanzadas técnicas basadas en redes neuronales y LSTM. Este enfoque se basó exclusivamente en los datos históricos de los precios de cierre de las 500 compañías que conforman el índice. La tarea consistió en clasificar los movimientos en tres categorías: 1 para comprar, 0

para mantenerse y -1 para vender (Agustín, 2017). La metodología utilizada para esta clasificación se fundamentó en las Bandas de Bollinger, las cuales establecen límites que indican si el precio se encuentra en niveles considerados demasiado altos o bajos. Como características para el análisis, se emplearon variables generadas a partir de los precios de cierre ajustados de las compañías del índice. Estas variables se derivaron considerando los componentes de un posible sistema de trading diseñado para generar señales de compra y venta (Agustín, 2017).

López y Sales, en el (2020) realizó su investigación el cual intitula “**Implementación hardware de una red neuronal Long Short-Term Memory**” en el marco de este Proyecto de Fin de Grado, el enfoque principal radica en la investigación y aplicación de redes neuronales profundas en hardware para la predicción de la glucemia a 30 minutos de anticipación. Específicamente, se optará por utilizar las redes neuronales LSTM, conocidas por proporcionar predicciones más precisas en comparación con otras técnicas clásicas, especialmente en el contexto de series temporales de datos. En el análisis de estos modelos, se consideran como entradas varios parámetros, tales como los niveles de glucemia, la presencia de hidratos de carbono e insulina en la sangre. En nuestro conjunto de datos, cada registro de características incluye un historial que se remonta a dos horas atrás hasta el momento actual (López & Sales, 2020). El presente trabajo comienza detallando la investigación llevada a cabo en Python para determinar la arquitectura de la red neuronal que será posteriormente implementada en hardware. Esta estructura abarca aspectos como el número de capas, el tipo de capas utilizadas y la cantidad de neuronas que las conforman (López & Sales, 2020).

Fierro, en el (2020) Realizó el trabajo para obtener el grado de especialista en Inteligencia de datos orientada al Big Data, la cual intitula “**Predicción de series temporales con redes neuronales**” donde el propósito fundamental de esta investigación es realizar una comparativa entre diversos enfoques de aprendizaje automático fundamentados en redes neuronales y el método estadístico vectorial autorregresivo en el ámbito de la predicción de series temporales multivariadas (Fierro, 2020), como conclusiones del trabajo se obtuvieron:

Es notable la sencillez computacional del modelo estadístico en comparación con las redes neuronales, un factor crucial al elegir un método predictivo para un problema

específico. El tiempo necesario para entrenar un solo modelo de aprendizaje automático es considerablemente menor. Además, se destaca que la diversidad de opciones para la preparación de datos y modelos dificulta una comparación exhaustiva o concluyente (Fierro, 2020).

El modelo estadístico muestra que sus patrones predictivos tienden a suavizarse o desaparecer a medida que se amplía el horizonte temporal. En contraste, los modelos de aprendizaje automático mantienen los patrones aprendidos, lo que podría explicar intuitivamente su mejor rendimiento en horizontes temporales más extensos. Además, las imágenes revelan que el modelo convolucional captura con mayor nitidez los patrones cíclicos en comparación con el modelo MLP (Fierro, 2020).

2.1.2. ANTECEDENTES NACIONALES

Saavedra, en (2017) publicó un artículo de investigación con el siguiente título **“Inversión pública en seguridad ciudadana y los indicadores de delitos, faltas y violencia familiar en el Perú, 2006 – 2016”** Su propósito fue identificar la conexión entre la inversión pública en seguridad ciudadana y los indicadores de delitos, infracciones y violencia familiar en Perú durante el periodo 2006-2016. La investigación adoptó un diseño no experimental, descriptivo, correlacional, con un enfoque longitudinal y una metodología cuantitativa (Saavedra, 2017).

Como resultado de nuestro análisis, se ha identificado un patrón en la ejecución presupuestaria durante el período comprendido entre 2006 y 2016 en el Perú. A nivel nacional, se observa un aumento progresivo y constante en la ejecución presupuestaria en ese lapso temporal, coincidiendo con un aumento correlativo en las denuncias por delitos. Además, se ha constatado que los delitos, las infracciones y los casos de violencia familiar, tanto de carácter físico como psicológico, han experimentado un incremento durante dicho período (Saavedra, 2017).

Dentro de este contexto, en la región de Madre de Dios, caracterizada por una ejecución presupuestaria relativamente baja, el promedio de inversión presupuestal entre los años 2009 y 2016 fue de S/. 2.304.425, mientras que el número promedio de denuncias por delitos alcanzó las 1325. Por otro lado, en la región de La Libertad, que destaca por una alta ejecución de gastos, el promedio de inversión presupuestaria fue de S/. 6.762.016, y el promedio de denuncias por delitos se situó en 10166. Estos hallazgos permiten concluir que, durante el período estudiado, se ha producido un aumento significativo en la incidencia de delitos, así como en los casos

de violencia familiar, tanto de carácter físico como psicológico, en ambas regiones. Concluyeron que “Los delitos denunciados muestran una tendencia creciente lo cual involucra un empeoramiento de la situación delictual a nivel nacional” (Saavedra, 2017)

(Rejas, Diaz, & Troya, 2022) realizó una publicación que titulaba **“La evaluación del presupuesto por resultados y su impacto en la seguridad ciudadana del Perú”**

El propósito de la investigación fue coordinar los esfuerzos de todas las entidades involucradas para abordar la implementación de la Estrategia Nacional de Seguridad Ciudadana 2018-2023. Este enfoque tiene repercusiones significativas en la distribución de recursos públicos esenciales destinados a satisfacer las demandas de la ciudadanía. El objetivo de la investigación era evaluar los resultados obtenidos a partir de los presupuestos asignados y destacar aspectos positivos y negativos relacionados con la gestión de los fondos públicos. Todo esto con la finalidad de contribuir a la mejora de las políticas públicas en este ámbito (Rejas, Diaz, & Troya, 2022).

El objetivo de la investigación era coordinar los esfuerzos de todas las entidades involucradas para abordar la necesidad de poner en marcha la Estrategia Nacional de Seguridad Ciudadana 2018-2023. Este enfoque tiene implicaciones significativas en la distribución de recursos públicos vitales destinados a satisfacer las demandas de la ciudadanía. La investigación buscaba evaluar los resultados obtenidos de los presupuestos asignados y detectar aspectos positivos y negativos vinculados con la gestión de los fondos públicos, con el fin de contribuir a la mejora de las políticas públicas en este ámbito (Rejas, Diaz, & Troya, 2022).

2.2. BASES TEÓRICAS

2.2.1. DENUNCIA

Según la Policía Nacional del Perú, denunciar consiste en comunicar a las autoridades los sucesos ocurridos para que tomen las medidas adecuadas. La denuncia representa la declaración que realiza una persona con el propósito de informar al Juez, al Ministerio Fiscal o a la policía sobre eventos que se sospecha pueden constituir un delito. Presentar una denuncia es simultáneamente un derecho y una responsabilidad de la víctima, con la finalidad de prevenir futuras situaciones

de vulnerabilidad. En el proceso de presentar la denuncia, es válido exigir un trato respetuoso y profesional (Defensoría del Pueblo, 2013).

Para el poder Judicial una denuncia es una manifestación, ya sea hablada o por escrito, mediante la cual se informa a la autoridad sobre cualquier hecho del cual se tenga conocimiento y que podría constituir una violación penal, incluso si no se presenció directamente ni causó perjuicio personal (Ministerio Público, 2024). Cualquier persona que sea testigo de la comisión de un delito público o que, aun sin presenciarlo, tenga conocimiento de él de alguna manera, está obligada a informarlo de inmediato a la autoridad, corriendo el riesgo de infringir la ley si no lo hace. Presentar una denuncia no implica responsabilidad ni obligación especial para el denunciante, ya que no está obligado a demostrar los hechos denunciados ni participar en el proceso judicial que pueda surgir como resultado de tales eventos (Ministerio público , 2024).

2.2.1.1. Requisitos para denunciar:

- Proporcionar información personal al incluir sus datos personales en la denuncia escrita o al mostrar documentos como la identificación, pasaporte, permiso de conducir u otros documentos similares.
- Describir los eventos que considere podrían constituir delitos, proporcionando todos los detalles disponibles sobre el incidente denunciado y las personas involucradas.
- Firmar la denuncia personalmente. En caso de no poder hacerlo, otra persona debe hacerlo en su nombre a su solicitud.

2.2.1.2. Procedimiento de la denuncia

El proceso de denuncia presenta como requisitos mínimos los siguientes aspectos:

1. La denuncia presentada ante la entidad debe ser registrada, reemplazando la identidad del denunciante por un código cifrado. La entidad procede a verificar si la denuncia cumple con los requisitos estipulados en el artículo 7. En caso contrario, se solicita al denunciante que corrija la omisión. Si no se cumple con esta corrección, la entidad asume que el denunciante ha renunciado a su solicitud y, por ende, a cualquier eventual petición de medidas de protección, lo que resulta en el archivo de la solicitud. No obstante, la entidad puede

evaluar la información proporcionada hasta ese momento y, si tiene fundamento, materialidad o interés propio, debe trasladarla a las instancias competentes (Ministerio público , 2024).

2. En caso de que la denuncia contenga elementos suficientes de importancia y fundamentos que justifiquen el inicio del procedimiento administrativo bajo la responsabilidad de la entidad, se procede a comunicar esta situación al titular de la Secretaría Técnica encargada de precalificar las presuntas faltas administrativas o a quien ejerza dicha función, de acuerdo con la normativa aplicable (Ministerio Público, 2024).
3. Si los hechos objeto de la denuncia se refieren a asuntos o disputas que están bajo la competencia de otros organismos estatales, se notifica al denunciante al respecto y se remite toda la documentación proporcionada a la entidad correspondiente, garantizando la confidencialidad de la misma (Ley N°30506, 2017).

2.2.1.3. Proceso de denuncia

Según la guía de denuncias del Consejo General del Poder Judicial, el curso que toman las denuncias varía según el lugar donde hayan sido presentadas. A continuación, se describen detalladamente los procedimientos correspondientes a cada caso.

- **Denuncias presentadas ante la autoridad policial:** En el caso de un delito público o cuando se requiera para prevenir la instrucción de diligencias por un motivo relacionado con un delito privado, se notificará de inmediato a la autoridad judicial o al representante del Ministerio. En general, este tipo de denuncias suele dar lugar a una investigación preliminar llevada a cabo por las autoridades policiales. El objetivo principal de esta investigación es verificar los hechos y recopilar declaraciones de testigos, así como otras diligencias que los funcionarios pertinentes consideren necesarias. En situaciones que involucren un delito público o cuando sea necesario evitar la instrucción de diligencias vinculadas a un delito privado, se realizará una notificación inmediata a la autoridad judicial o al representante del Ministerio (Ministerio público , 2024).

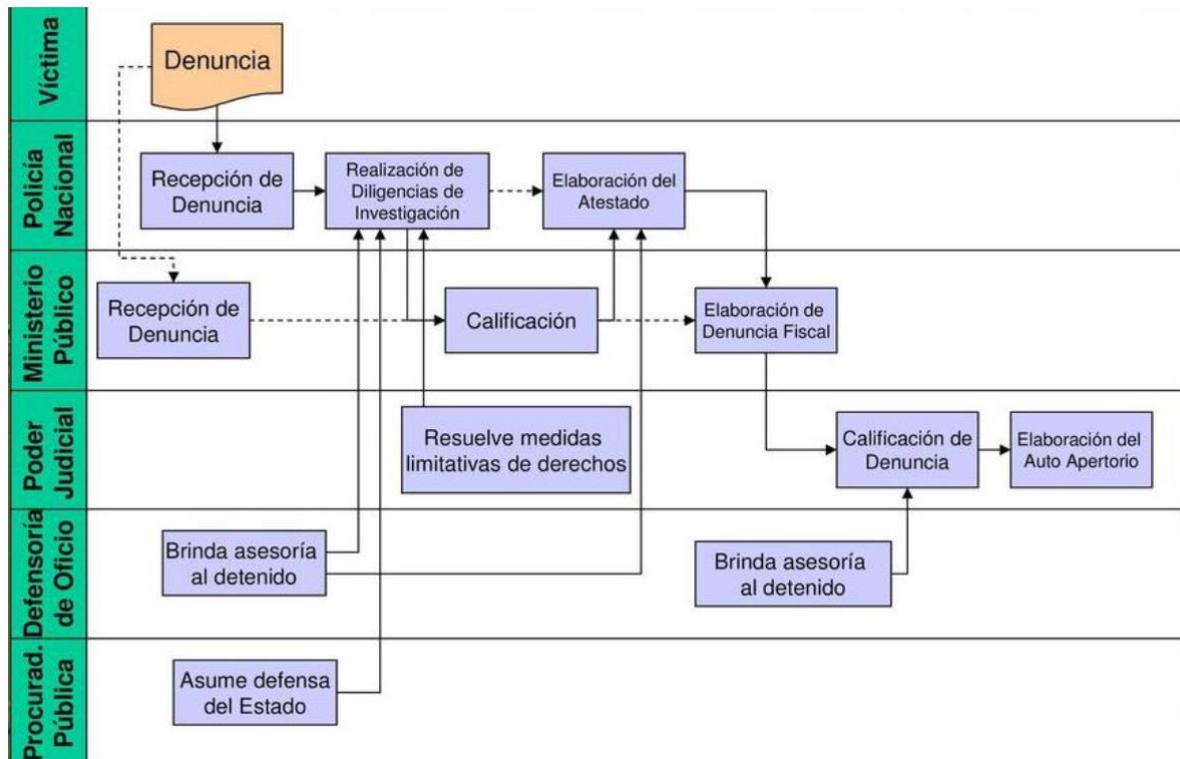
- **Denuncias presentadas ante el Ministerio Fiscal:** Este tipo de denuncias suele implicar que el Ministerio realice directamente algunas verificaciones o que encargue a las fuerzas policiales la realización de una investigación preliminar sobre los hechos denunciados. Después de esta investigación, si el Fiscal determina que hay indicios de una posible infracción penal, remitirá toda la documentación al tribunal competente (Ministerio público , 2024).
- **Denuncias presentadas ante el Juzgado de Instrucción:**

Al llegar al tribunal, el juez revisa el contenido de la denuncia y determina las acciones a seguir (Ministerio público , 2024). Después de llevar a cabo las diligencias de investigación pertinentes, puede tomar las siguientes decisiones:

 - a) Archivar el caso: si considera que los hechos denunciados no constituyen una violación penal.
 - b) Sobreseimiento provisional: si, a pesar de que los hechos denunciados parecen constituir una infracción penal, no se ha logrado identificar o localizar al presunto responsable, hasta que este sea encontrado.
 - c) Trasladar el expediente a:
 - El tribunal del área donde ocurrió el hecho denunciado, si es diferente al tribunal que recibió la denuncia.
 - El tribunal de menores si los posibles responsables son menores.
 - El tribunal militar si el hecho está dentro de la competencia de la jurisdicción militar.
 - d) Iniciar un procedimiento penal, el cual puede variar según la gravedad y características del hecho denunciado, si el juez considera que inicialmente constituye una infracción penal atribuible a una o varias personas identificadas.

Figura 1

Estructura del proceso de una denuncia



Nota. Tomado de la tesis: *Aplicación móvil para la gestión de denuncias de los efectivos policiales en la Policía Nacional del Perú*, Ccopa, (2019)

2.2.2. DENUNCIANTE

Hace referencia a cualquier persona o entidad, ya sea de forma individual o colectiva, que notifica a la institución a través de sus órganos competentes acerca de un acto de corrupción o delito. El denunciante desempeña el papel de colaborador externo tanto de la Administración Pública como del Sistema de Justicia. No obstante, no se le clasifica como parte integrante del proceso administrativo disciplinario, del procedimiento administrativo funcional o del proceso penal que pueda originarse como resultado de su denuncia (Ccopa, 2019).

La nacionalidad, sexo, minoría de edad, residencia, incapacidad legal, internamiento en un centro de readaptación social o reclusión, asistencia escolar, hospitalización, clínica u otras formas de sujeción o dependencia del denunciante hacia una tercera persona no constituyen obstáculos para presentar una denuncia (Ccopa, 2019).

2.2.3. DENUNCIADO

Persona o entidad que es objeto de una denuncia. Es aquella contra la cual se presenta formalmente un señalamiento o acusación, generalmente relacionada con la comisión de un acto ilícito, irregularidad o conducta indebida. En el contexto legal y jurídico, el denunciado es el individuo o entidad que se presume ha participado en un hecho que se considera contrario a la ley o normativa, y que está siendo señalado como responsable de dicha acción (Ccopa, 2019).

2.3. REDES NEURONALES ARTIFICIALES

2.3.1. REDES NEURONALES BIOLÓGICAS

Una Red Neuronal Artificial (RNA) representa un modelo matemático que se inspira en el funcionamiento biológico de las neuronas y en la organización estructural del cerebro humano. Desde una perspectiva más amplia, puede concebirse como un sistema inteligente que ejecuta operaciones de manera divergente en comparación con las computadoras convencionales (Bertona, 2005). Aunque las computadoras más avanzadas de la actualidad destacan por su velocidad en la manipulación de información, se encuentran limitadas en la realización de tareas de gran complejidad, como el reconocimiento y la clasificación de patrones. Estas tareas continúan requiriendo un considerable gasto de tiempo y recursos computacionales, incluso en los sistemas más poderosos disponibles en la actualidad. No obstante, el cerebro humano exhibe una capacidad superior para abordar estas tareas, a menudo con aparente facilidad. Un ejemplo ilustrativo es la habilidad para reconocer un rostro familiar en medio de una multitud de otros rostros (Cabello, 2004).

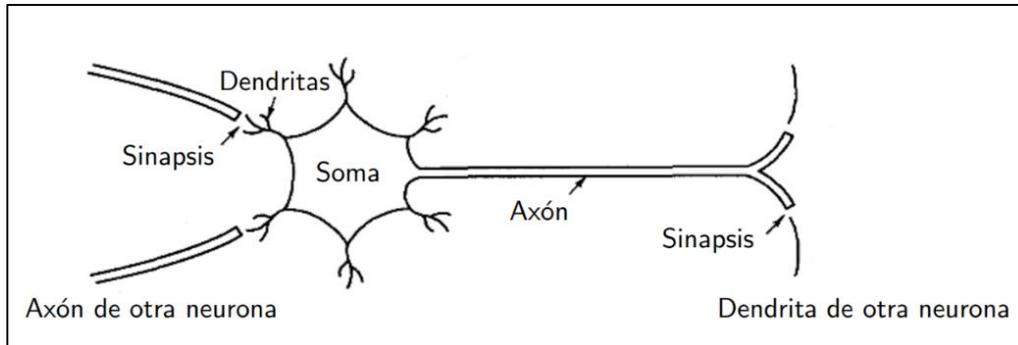
El cerebro se puede concebir como un sistema de una complejidad excepcional. Su unidad elemental, la neurona biológica, se encuentra distribuida de manera masiva, estableciendo interconexiones significativas entre ellas. Se estima que, en la corteza cerebral, por ejemplo, se alojan aproximadamente 10 billones de neuronas, las cuales establecen asombrosamente alrededor de 60 trillones de conexiones sinápticas. Aunque existen diversos tipos de neuronas biológicas, la Figura 2 presenta una representación simplificada de un tipo específico que es ampliamente prevalente (Tablada, sf). Observamos que esta entidad neuronal se compone de:

- El soma, identificado como el cuerpo central, que alberga el núcleo celular.
- El axón, una prolongación del soma.

- Las dendritas, ramificaciones terminales que forman parte de la estructura.
- La sinapsis, una zona de conexión entre una neurona y otra.

Figura 2

Modelo simplificado de una neurona biológica



Nota. La Imagen fue tomada de *Redes Neuronales Artificiales* (Tablada, sf).

2.2.1.1. Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) representan una aproximación que, en ciertos aspectos, emula la estructura y funcionalidad del sistema nervioso, dando lugar a sistemas de procesamiento altamente paralelos, distribuidos y adaptables, que exhiben características de comportamiento que podríamos calificar como "inteligentes". Sin embargo, al profundizar en el análisis de los algoritmos y comportamientos cerebrales, se hace evidente que las diferencias entre una computadora y el cerebro son considerablemente más marcadas de lo que podría suponerse inicialmente (Muñoz, 1996).

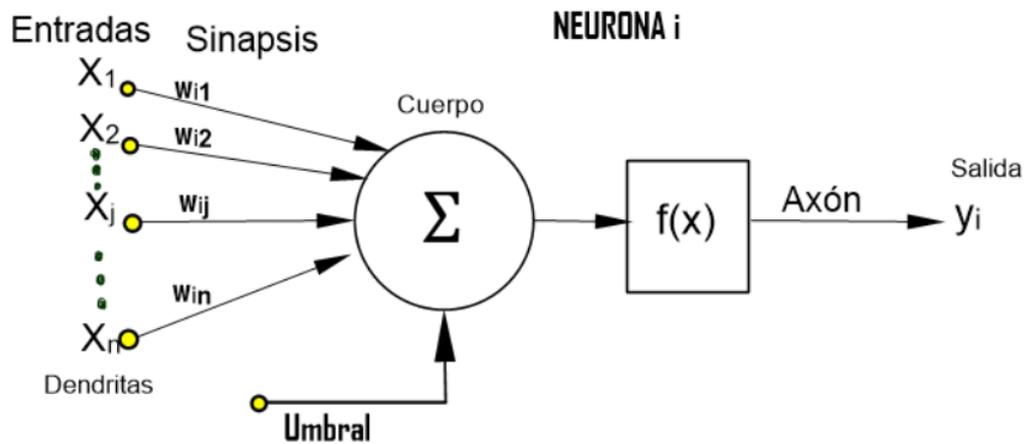
El elemento fundamental en las RNA es la neurona artificial, la cual se organiza en diversas capas que se interconectan para formar una red neuronal. Estas redes neuronales, ya sea individualmente o interconectadas entre sí, tienen la capacidad de constituir un sistema global de procesamiento (Namigtle, 2018) .

2.2.1.2. Elementos de una Red Neuronal Artificial

Se presenta en la Figura 3 un modelo elemental de una neurona artificial, resaltando su similitud con la neurona biológica (Matich, 2001).

Figura 3

Modelo de neurona estándar



Nota. Imagen tomada del libro *Redes Neuronales: Conceptos Básicos y Aplicaciones* (Matich, 2001)

Una red neuronal tiene las siguientes partes:

a. Neurona

Una Red Neuronal Artificial (RNA), muchas veces conocida como red neuronal, se compone mediante la interconexión de múltiples unidades de procesamiento denominadas neuronas, las cuales son bloques no lineales distribuidos a lo largo y ancho de toda la red neuronal (Acosta-Cervantes, Villareal-Marroquin, & Cabrera-Rios, 2013).

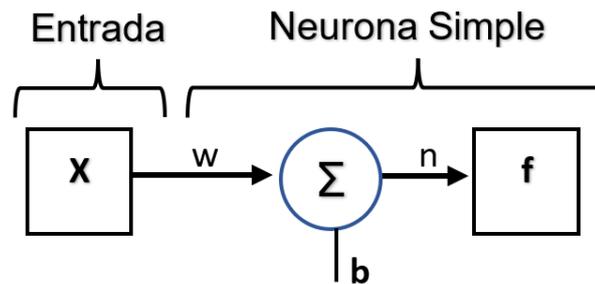
Por lo general, las neuronas se organizan en capas, y una o varias de estas capas conforman la estructura global de la red neuronal. En las redes neuronales, es posible distinguir tres tipos de capas: la capa de entrada, la capa oculta y la capa de salida. La capa de entrada, también denominada capa sensorial, se encuentra compuesta por neuronas que asumen la responsabilidad de recibir datos o señales procedentes del entorno, como, información proveniente de sensores o retroalimentación derivada de las salidas previas. La capa de salida alberga neuronas que proporcionan la salida final o la respuesta de la red. Por su parte, la capa oculta, o capa intermedia, no mantiene conexiones directas con el entorno. Su función radica en ampliar los grados de libertad de la red,

enriqueciendo, de esta manera, su capacidad computacional (Acosta-Cervantes, Villareal-Marroquin, & Cabrera-Rios, 2013).

Las redes neuronales se construyen interconectando neuronas, se muestra el bloque principal de una neurona de entrada única.

Figura 4

Esquema de una Neurona Simple



Fuente: imagen capturada del trabajo Redes Neuronales Recurrentes: Una aplicación para los mercados bursátiles (Agustín, 2017)

x: entrada escalar

w: peso escalar

f: función de transferencia

n: entrada neta: Net

La salida está dada por la siguiente ecuación:

$$y = f(n) = f(w \cdot x + b) \quad (1)$$

b. Pesos

Llamados factores de relevancia o parámetros, en el contexto de cada neurona, se les asigna un peso específico. Estos pesos son valores numéricos que experimentan modificaciones a lo largo del proceso de aprendizaje de la red neuronal. Los pesos actúan como portadores de información crucial que determinará la funcionalidad específica de la red neuronal. En particular, se emplea la notación "w" para representar los pesos asociados a las capas ocultas, mientras que la letra griega β se utiliza para denotar los pesos relacionados con la capa de salida. (Basoqain, sf).

c. Regla de Propagación

Llamada también unión sumadora. Es la operación que permite sumar las entradas teniendo en consideración la importancia o peso de cada una de las mismas. Esta combinación de entradas con pesos es el Net (Fernandez, 2021).

$$Net_m = w_m^T X + w_0; Net_k = \beta_k^T Z + \beta_0$$

Donde: $w_m^T = (w_{1m}, w_{2m}, \dots, w_{pm})$ y $\beta_k^T = (\beta_{1k}, \beta_{2k}, \dots, \beta_{Mk})$

d. Función de Activación o Función de Transferencia

El valor que se obtiene de la unión sumadora o regla de propagación (Net, se filtra mediante el uso de la función de activación. Esta función genera el valor final de la salida de la red.

A través de la siguiente ecuación, una función de transferencia procesa la entrada, que puede variar entre menos infinito y más infinito, para producir una salida limitada entre 0 y 1:

$$y = \frac{1}{1 + e^{-n}}$$

La función logarítmica sigmoide se emplea comúnmente en redes multicapa entrenadas con algoritmos de retropropagación, ya que se destaca por su capacidad de ser diferenciable.

En la Tabla 1 se muestran las funciones de transferencia más utilizadas

Tabla 1

Funciones de Transferencia

Nombre	Relación entre entrada y salida
Hard limit Limitador fuerte	$Y = 0, \quad si \ n < 0$ $Y = 1, \quad si \ n \geq 0$
Limitador fuerte simétrico	$Y = -1, \quad si \ n < 0$ $Y = 1, \quad si \ n \geq 0$
Lineal	$Y = n$
Lineal Saturada	$Y = -1, \quad si \ n < -1$ $Y = n, \quad si \ -1 < n < 1$ $Y = 1, \quad si \ n \geq 1$
Sigmoide logarítmica	$Y = \frac{1}{1 + e^{-n}}$
Sigmoide tangente hiperbólica	$Y = \frac{e^n - e^{-n}}{e^n + e^{-n}}$
Competitiva	$Y = 1$ para neuronas con n máximo $Y=0$ otros casos

Nota. Tomado de *Redes Neuronales: Conceptos Básicos y Aplicaciones* (Matich, 2001)

e. Salida

Es el resultado de la función de transferencia:

$$f(X) = g(Net)$$

2.2.1.3. Tipos de Redes Neuronales Artificial

Las redes neuronales pueden ser utilizados además como métodos estadísticos para dar solución a problemas de clasificación o regresión.

a. Regresión.

Dado un v.a. (vector aleatorio) $X \in \mathbb{R}^p$ y la variable respuesta $Y \in \mathbb{R}$ el modelo de regresión está definido por:

$y = f(X) + \varepsilon$ El propósito es seleccionar f para predecir o estimar Y , dado algún valor de X .

Para el ajuste del modelo se compara la salida real con el valor obtenido por la red $f(X)$, mediante una medida, conocida como función de pérdida

$L(Y, f(X))$ que penaliza o sanciona el error de predicción (Agustín, 2017). La función de pérdida para la regresión es comúnmente el error cuadrático:

$$L(Y, f(X)) = (Y - f(X))^2$$

Para estimar f , previamente definimos el error esperado de predicción, como la esperanza de $L(Y, f(X))$, esto es:

$$EPE(f) = E[L(Y, f(X))] = E[(Y - f(X))^2]$$

$$= \int (y - f(x))^2 P(dx, dy)$$

$$= \int (y - f(x))^2 P(dy/dy)P(dx)$$

$$EPE(f) = E_x E_{\frac{Y}{X}}([Y - f(X)]^2 / X)$$

$$f(x) = \operatorname{argmin}_c E_{y/X}([Y - c]^2 / X = x)$$

la solución es:

$$f(x) = E(y/X = x)$$

b. Clasificación

Cuando nos enfrentamos a un problema de clasificación, la variable de salida, denotada como G en lugar de Y , representa una categoría categórica $G \in C$. En este contexto, G toma valores discretos en un conjunto finito $C = \{1, 2, \dots, K\}$,

donde estos números actúan como etiquetas que representan clases $\{G_1, G_2, \dots, G_K\}$.

Para llevar a cabo la clasificación, se emplea una regla designada como $G(X)$, donde $G(X) = k$ indica que un objeto con un vector de características es asignado a la k -ésima clase G_k . Específicamente, esta regla tiene como objetivo la construcción de una partición del espacio de características en K regiones mutuamente excluyentes, identificadas como R_1, \dots, R_K .

En el contexto de la clasificación, se necesita una función de pérdida diferente a la definida en el caso de la regresión. La función de pérdida más comúnmente utilizada para la clasificación es la función 0-1, cuya definición es la siguiente:

$$L(G, G(X)) = I(G \neq G(X)) = \begin{cases} 1 & G = G(X) \\ 0 & G \neq G(X) \end{cases}$$

Y el valor esperado del error de predicción mediante:

$$EPE = E[L(G, G(X))]$$

donde la esperanza es tomada con respecto a la distribución conjunta $P(G, X)$

$$EPE = \sum_{k=1}^K L(G_k, G(X))P(G_k, X)P(X) = E_X \sum_{k=1}^K L(G_k, G(X))P(G_k, X)$$

minimizando la esperanza del valor de predicción se estima el clasificador

$$\hat{G}(X) = \operatorname{argmin}_{g \in C} \sum_{k=1}^k L(G = k, g)P(G = g/X = x)$$

simplificando como la función pérdida asume valores 0 o 1 se tiene:

$$\hat{G}(X) = \operatorname{argmin}_{g \in C} [1 - P(G = g/X = x)]$$

$$\hat{G}(X) = \operatorname{argmin}_{g \in C} [P(G = g/X = x)]$$

$$\hat{G}(X) = k \quad \text{si} \quad P(G = g/X = x) = \max_{g \in C} P(G = g/X = x)$$

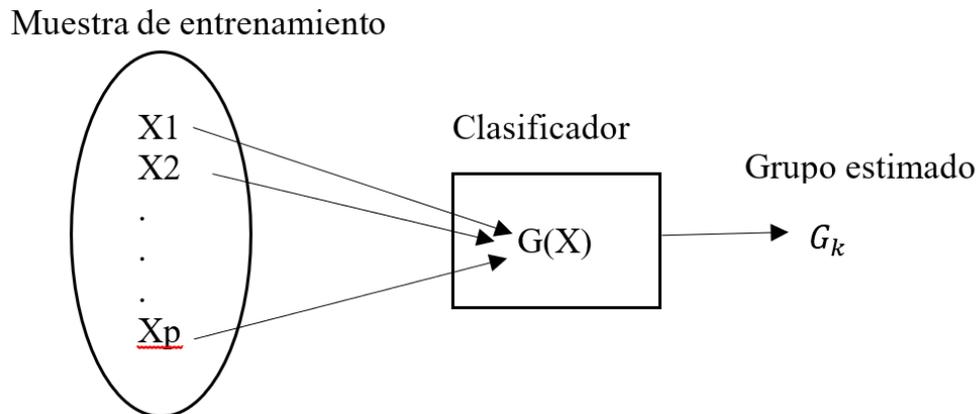
A esta solución se le denomina clasificador de Bayes.

En Resumen, el problema de clasificación consta de los siguientes pasos

- Dado un conjunto de entrenamiento E de N observaciones, cada observación X_i en E pertenece a una de las K clases $\{G_1, \dots, G_K\}$.
- Con el conjunto de E , se construye una función $G(X): R^p \rightarrow \{G_1, \dots, G_k\}$ a esta función se le denomina clasificador.
- El clasificador $G(X)$ asigna a la observación X , a una de las clases $G(X)$ o $\{G_1, \dots, G_k\}$

Figura 5

Representación de un Clasificador



Nota. Imagen extraída de la tesis *Implementación hardware de una red neuronal Long Short-Term Memory* (López & Sales, 2020)

2.2.1.4. Perceptrón

En este modelo, se realiza una suma ponderada de las señales de entrada y se compara con un umbral o límite para determinar la salida de la neurona. Si la suma es mayor o igual al umbral, la salida es 1; en caso contrario, la salida es 0. Las redes que incorporan estas neuronas tienen la capacidad de calcular diversas operaciones aritméticas o funciones lógicas. A diferencia de las redes biológicas, los parámetros de estas redes simples deben diseñarse, ya que no existe un método de entrenamiento preestablecido. No obstante, este enfoque ha generado un considerable interés y ha inspirado el desarrollo de computadoras digitales basadas en la estructura de las neuronas biológicas (Maya, 2022).

Las redes neuronales iniciales, denominadas perceptrones, se asemejaban en su arquitectura a las redes concebidas por McCulloch y Pitts. Sin embargo, se introdujo una innovadora regla de aprendizaje con el propósito de instruir a los perceptrones en la resolución de problemas de reconocimiento de patrones. Este avance se erigió como un hito fundamental en la progresión de las redes neuronales, dado que se demostró que su regla de aprendizaje inevitablemente converge hacia la correcta ponderación de los parámetros de la red, siempre y cuando existan configuraciones de pesos susceptibles de resolver el problema planteado.

Se presentaron ejemplos de comportamiento deseado para la red, permitiendo que esta aprendiera de sus propios errores gracias a la regla de aprendizaje. Esta

metodología permitió que el perceptrón adquiriera conocimiento, incluso cuando se inicializaba con valores aleatorios para los pesos y sesgos. No obstante, los perceptrones en su forma original exhibían limitaciones notables, siendo incapaces de implementar algunas funciones elementales. Estas restricciones, sin embargo, se superaron en la década de 1980 con la introducción de perceptrones multicapa mejorados y las correspondientes reglas de aprendizaje.

En la actualidad, el perceptrón sigue siendo reconocido como una red importante, siendo una opción rápida y fiable para abordar aplicaciones simples. Explorar el funcionamiento del perceptrón también sienta las bases para comprender redes más complejas.

La operación de una neurona biológica real implica la emisión de una señal de salida solo cuando la suma total de las señales de entrada supera un umbral establecido. Este principio se replica en el perceptrón, donde se calcula la suma ponderada de las entradas que representan todas las señales de entrada. La salida de esta operación alimenta otras redes perceptrones, imitando así el comportamiento de las redes neuronales biológicas.

Considerando el vector aleatorio $X \in R^p$ y la variable dependiente $Y \in R$, el perceptrón se configura como una red neuronal con una neurona en la capa oculta, definida por la relación:

$$y = f(X) + \varepsilon = g(w_0 + w^T X) + \varepsilon, \quad Net = w_0 + w^T x \quad (23)$$

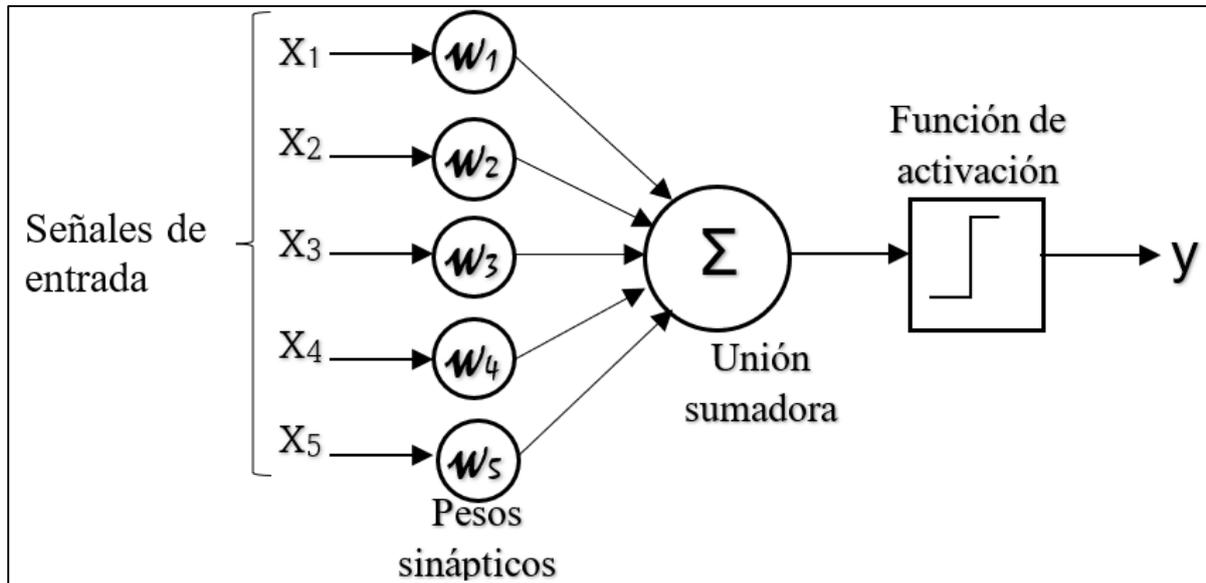
Este modelo representa la red neuronal más elemental, compuesta por una combinación lineal que incluye varias neuronas para recibir las entradas a la red y una neurona de salida. Su aplicación se restringe a problemas linealmente separables, actuando como un clasificador lineal definido por la función:

$$g(Net) = \begin{cases} 0 & \text{cuando } Net < 0 \\ 1 & \text{cuando } Net \geq 0 \end{cases} \quad (24)$$

Cuando se tiene un conjunto de entrenamiento $X \in R^p$, el objetivo del perceptrón es clasificarlo en uno de dos grupos, ya sea G_1 o G_2 . La regla de decisión para esta clasificación consiste en asignar x al grupo G_1 si la salida del perceptrón es 1, y al grupo G_2 si es 0.

Figura 6

Representación de un perceptrón



Nota: Imagen tomada de la tesis *Implementación hardware de una red neuronal Long Short-Term Memory* (López & Sales, 2020)

2.2.1.5. Aprendizaje del Perceptrón

El algoritmo de aprendizaje del perceptrón se esfuerza por descubrir una separación en el espacio de entrada, buscando minimizar el error cuadrático medio, como se define por:

$$R = \sum_{i=1}^N L(y_i, f(x_i)) = \frac{1}{2} \sum_{i=1}^N (y_i - f(x_i))^2 = \sum_{i=1}^N R_i \quad (25)$$

donde $f(x) = g(Net)$

La función de error R, utilizada en el conjunto de entrenamiento, se elige principalmente por su simplicidad analítica, pero puede tener diversas expresiones según la aplicación específica.

Esta función, denotada por R, es una expresión matemática definida en R^p , creando una superficie que presenta numerosos mínimos locales y un mínimo global para un conjunto de entrenamiento dado.

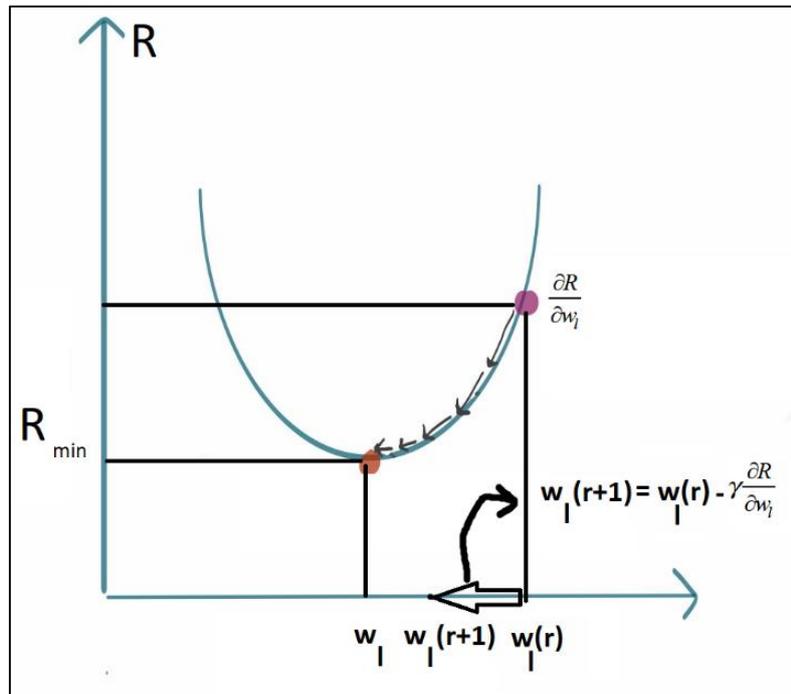
El proceso de actualización de los pesos implica realizar pequeños desplazamientos en el espacio de los pesos, siguiendo la dirección donde la función de error disminuye con mayor rapidez. Este procedimiento, conocido como método de

gradiente descendente, implica la modificación de los pesos y la obtención de información local sobre la superficie.

La aplicación del gradiente a la función de error, en un punto específico, determina la dirección de crecimiento máximo.

Figura 7

Ilustración del método de gradiente descendente



Nota. Imagen tomada de tesis *Combinación de clasificadores en redes neurales* (De la Torre, 2007)

La variación de los pesos guarda proporción con el descenso del gradiente de la función de error.

$$\Delta w_l = -\gamma \frac{\partial R}{\partial w_l} \implies -\gamma \sum_{i=1}^N \frac{\partial R}{\partial w_l} \quad l = 1, \dots, p, \quad \Delta w_l = w_l(r+1) - w_l$$

Utilizando la regla de la cadena para el realizar el cálculo de $\frac{\partial R}{\partial w_l}$

$$\frac{\partial R_i}{\partial w_l} = \frac{\partial R_i}{\partial \text{Net}} \frac{\partial \text{Net}}{\partial w_l}$$

y usando la función de transferencia identidad $g(\text{Net}) = \text{Net}$, se tiene:

$$\frac{\partial R_i}{\partial \text{Net}} = \frac{\partial}{\partial \text{Net}} \left[\frac{1}{2} (y_i - f(x_i))^2 \right] = -(y_i - f(x_i)) = -R_i$$

y

$$\frac{\partial \text{Net}}{\partial w_l} = \frac{\partial \sum_{i=1}^p w_l x_{il}}{\partial w_l} = x_{il}$$

Reemplazando se tiene:

$$\frac{\partial R_i}{\partial w_l} = -R_i x_{il}$$

La actualización de los pesos en la r-ésima iteración para el conjunto de entrenamiento es:

$$w_l(r+1) = w_l(r) + \gamma \sum_{i=1}^N R_i x_{il}$$

Donde:

- γ se denomina factor de aprendizaje el cual varía entre 0 y 1

El objetivo del aprendizaje del perceptrón es hallar un separador en forma de una línea del espacio de entrada.

2.2.1.6. Red Neuronal Multicapa

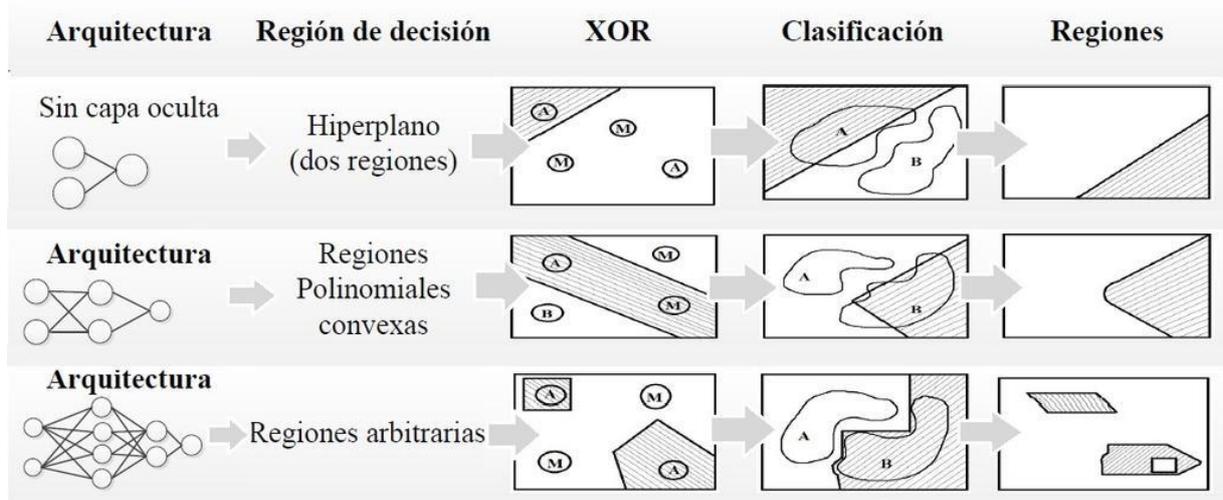
La capacidad del perceptrón se limita a resolver problemas que sean linealmente separables, los cuales son casos particulares y simplificados en comparación con los problemas del mundo real. Ante esta limitación, surge la pregunta de si es factible combinar elementos de procesamiento tipo perceptrón en múltiples capas, y en caso afirmativo, cómo se podría entrenar tal red.

La idea detrás de esta combinación de 'perceptrones' en varias capas sería dividir el espacio de las entradas en subespacios, cada uno representando una característica específica del problema. Luego, en capas subsiguientes, se detallarían de manera más específica las distintas regiones del espacio de características, hasta lograr la salida deseada, es decir, la partición del espacio de entradas.

Al disponer los perceptrones en cascada mediante una estructura de múltiples capas, se puede implementar fronteras de decisión mucho más complejas, como se muestra en la figura 8.

Figura 8

Regiones de Decisión



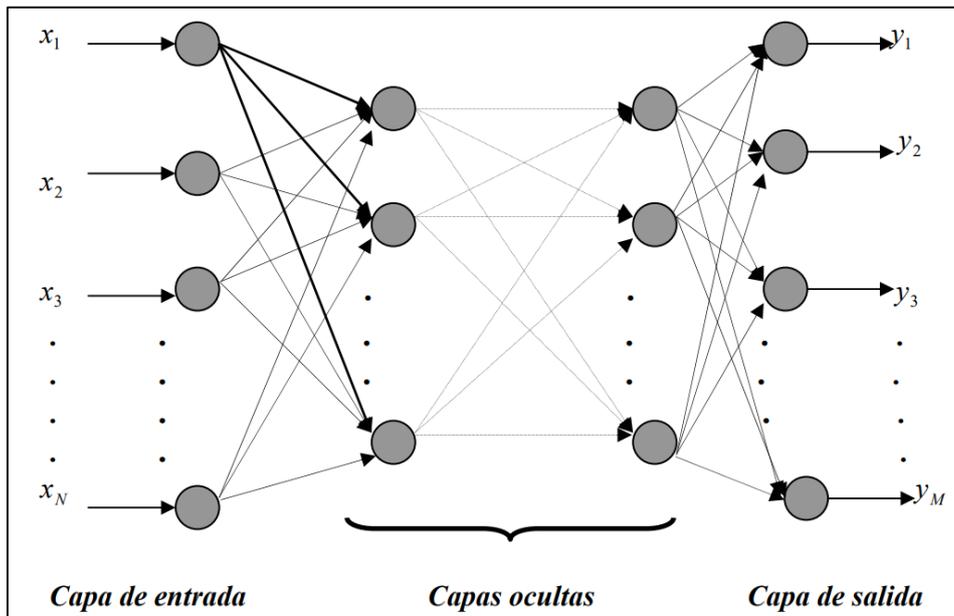
Fuente: Imagen extraída del libro *Redes Neuronales y Sistemas Borrosos* (Martin del Brio & Sanz, 2006)

2.2.1.7. Estructura de una Red Neuronal Multicapa

La estructura común de una red multicapa se observa en la siguiente figura

Figura 9

Representación de una Red Neuronal Multicapa



Nota: Imagen tomada de la tesis *Estudio y predicción de activos financieros mediante redes neuronales* (Santamaría, 2019).

La primera capa es llamada capa de entrada que toma datos del conjunto de entrenamiento E y la última capa es denominada de salida, cuyos pesos adquieren, mediante un proceso de aprendizaje, distintos valores para observaciones distintas.

2.2.1.8. Algoritmo de propagación hacia atrás o Backpropagation

El algoritmo de propagación hacia atrás, en inglés backpropagation (BP), es una generalización del algoritmo de mínimos cuadrados parciales: LMS (del inglés, Least-Mean-Square algorithm). Se utiliza para entrenar redes multicapa. Al igual que con la ley de aprendizaje de LMS, el algoritmo BP utiliza método de gradientes descendentes, en el que el índice de rendimiento es el error cuadrático medio (Vorobioff, Cerrotta, Morel, & Amadio, 2022)

El algoritmo LMS y el de propagación hacia atrás difieren en la forma que calculan las derivadas. En las redes multicapa que utilizan funciones de transferencia no lineales, resulta complejo determinar la relación entre los pesos de la red y el error. Para calcular las derivadas, necesitamos usar la regla de la cadena. Este algoritmo de propagación hacia atrás (Backpropagation) es muy utilizado, también es conocido como algoritmo de retro propagación, debido a que el error se propaga en dirección inversa al cálculo de las salidas (UNIR, 2024).

La importancia de este proceso consiste en que a medida que se entrena la red las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada. Varias investigaciones han demostrado que durante el proceso de entrenamiento la red Backpropagation tiende a desarrollar relaciones internas entre neuronas con el fin de organizar los datos de entrenamiento en clases (Vorobioff, Cerrotta, Morel, & Amadio, 2022)

2.2.1.9. Regla de aprendizaje de Redes Multicapa.

El proceso de aprendizaje para entrenar una red neuronal multicapa implica varios pasos. En primer lugar, se debe definir la topología de la red, lo que incluye determinar el número de neuronas en la capa de entrada, que depende de la cantidad de componentes del vector de entrada. También se deben especificar la cantidad de capas ocultas, el número de neuronas en cada capa oculta, y finalmente, el número

de neuronas en la capa de salida, que se relaciona con el número de componentes del vector de salida o patrones objetivo (Vorobioff, Cerrotta, Morel, & Amadio, 2022). Una vez definida la topología, se asignan valores iniciales a los pesos (parámetros) que conforman la red. Estos pesos son cruciales ya que determinan la contribución de cada entrada a la salida de la red. Este proceso de asignación de pesos iniciales es fundamental para iniciar el proceso de entrenamiento de la red neuronal.

Es crucial destacar que no hay una técnica definitiva para determinar el número de capas ocultas ni la cantidad de neuronas que debería contener cada una de ellas para abordar un problema específico. En cambio, esta elección se basa en la experiencia del experimentador.

Durante el proceso de entrenamiento, cada observación de entrenamiento se propaga a través de la red, generando una respuesta en la capa de salida. Esta respuesta se compara con la salida real para calcular el error en el aprendizaje. Minimizar este error en cada iteración del proceso de aprendizaje indica la dirección adecuada para actualizar los pesos. Estos ajustes en los pesos, al final del entrenamiento, conducirán a una respuesta satisfactoria para todas las observaciones de entrenamiento.

Hemos desarrollado este procedimiento considerando una red con una capa de entrada, una capa oculta y una capa de salida. La salida de la red está determinada por:

$$f_k(X) = g[\beta_{0k} + \beta_k^T \sigma(w_{0m} + w_m^T X)] = g(Net_k)$$

Para comenzar el entrenamiento se le otorga a la red una observación x_i del conjunto de entrenamiento $E = \{(x_i, y_i) / i = 1, \dots, N\}$ la cual tiene p componentes.

$$x_i^T = (x_{i1}, \dots, x_{il}, \dots, x_{ip})$$

cuando se le presenta a la red una observación de entrenamiento, ésta se propaga a través de las conexiones existentes produciendo una Net en cada una de las M neuronas de la capa oculta mediante:

$$Net_{mi} = w_m^T x_i + w_{0m} = \sum_{l=1}^p w_{ml} x_{il} + \alpha_{0m},$$

$$m = 1, \dots, M \quad i = 1, \dots, N$$

$$w_m^T = (w_{m1}, \dots, w_{ml}, \dots, w_{mp})$$

Donde

- w_{ml} : peso que corresponde al componente l de la entrada con la neurona m de la capa oculta
- w_{0m} : umbral de la neurona m de la capa oculta

La salida de cada neurona en la capa oculta es Z_{mi} .

$$Z_{mi} = \sigma(Net_{mi}) = \sigma\left(\sum_{l=1}^p w_{ml}x_{il} + w_{0m}\right)$$

Donde

- σ : función de transferencia de las neuronas de la capa oculta

Las salidas Z_{mi} de las neuronas de la capa oculta (de M componentes) son las entradas a los pesos de conexión de la capa de salida para la i -ésima observación del conjunto de entrenamiento, este comportamiento está descrito por la siguiente ecuación

$$Net_{ki} = \beta_{0k} + \beta_k^T Z_i = \sum_{m=1}^M \beta_{mk} Z_{mi} + \beta_{0k}, \quad \beta_k = (\beta_{1m}, \dots, \beta_{km}, \dots, \beta_{kM})$$

Donde

- β_{km} : El peso asociado a la neurona m en la capa oculta con respecto a la neurona k de la capa de salida, la cual cuenta con M neuronas.
- Z_{mi} : salida de la neurona m de la capa oculta, la cual cuenta con M neuronas.
- β_{0k} : ganancia de la neurona k de la capa de salida.
- Net_{ki} : entrada neta a la neurona k de la capa de salida

La red neuronal genera una salida final para cada elemento de entrada como se puede visualizar en la siguiente ecuación

$$f_k(x_i) = g(Net_{ki}) \quad k = 1, \dots, K$$

donde

g : función de transferencia de las neuronas de la capa de salida y K es el número de neuronas en la capa de salida.

Reemplazando se obtiene la salida de la red en función de la entrada neta (Net_{ki}) y de los pesos de conexión con la última capa oculta.

$$f_k(x_i) = g(Net_{ki}) = g(\beta_{0k} + \beta_k^T Z) = g\left(\beta_{0k} + \beta_k^T \sigma(w_{0m} + w_m^T x)\right)$$

La salida de la red neuronal $f_k(x_i)$ se contrasta con el valor real y_{ik} para calcular una medida de error llamado función de pérdida $L(y_{ik}, f_k(x_i))$ para penalizar el error de predicción (Gulli & Pal, 2017). Cuando la salida y_{ik} es una variable aleatoria continua, la función de pérdida comúnmente está dada por el error cuadrático:

$$L(y_{ik}, f_k(x_i)) = \sum_{k=1}^K (y_{ik} - f_k(x_i))^2$$

Cuando la salida G es una variable categórica, se requiere de otra función de pérdida denominada cross-entropía o deviance definida por:

$$L(G_{ik}, f_k(x_i)) = - \sum_{k=1}^K I(G_{ik} = k) \log f_k(x_i) = - \sum_{k=1}^K G_{ik} \log f_k(x_i)$$

Donde $f_k(x)$: representa probabilidades, $f_k(x) = p_k(x)$ y está dado por la salida de la red.

La entropía evalúa la cantidad de información presente en el proceso de aprendizaje. La optimización de una red neuronal multicapa busca reducir al mínimo la función de error; $\sigma = \operatorname{argmin} R(\theta)$, donde

$$R(\theta) = \sum_{i=1}^N L(y_{ik}, f_k(x_i)) \text{ y } \theta = (w_{ml}, w_{0m}/m = 1, \dots, M; \beta_{km}, \beta_{k0}/k = 1, \dots, K)$$

El método de gradiente descendiente se emplea para obtener los valores de los pesos que minimizan la función de error.

2.2.1.10. Método de Gradiente Descendiente en Red Multicapa

Se emplea el espacio de pesos para minimizar el error durante el proceso de aprendizaje, el cual conforma una superficie de error con mínimos locales y un mínimo global (Acosta-Cervantes, Villareal-Marroquin, & Cabrera-Rios, 2013). Al analizar el gradiente del error en un punto de esta superficie, se identifica la dirección con el mayor crecimiento de la función de error. Con el objetivo de disminuir el error, se sigue la dirección negativa del gradiente para obtener el mayor decremento del error y lograr su minimización (Martín del Brio & Sanz, 2006). En la r -ésima iteración, las ecuaciones que gobiernan la actualización de los pesos para la capa oculta y la capa de salida son, respectivamente:

$$w_{ml}^{(r+1)} = w_{ml}^r - \gamma \sum_{i=1}^N \frac{\partial R_i}{\partial w_{ml}}$$

$$\beta_{km}^{(r+1)} = \beta_{km}^r - \gamma \sum_{i=1}^N \frac{\partial R_i}{\partial \beta_{km}}$$

donde

- γ : es el factor de aprendizaje que administra la velocidad del aprendizaje

Consideraciones en el algoritmo de Gradiente Descendiente

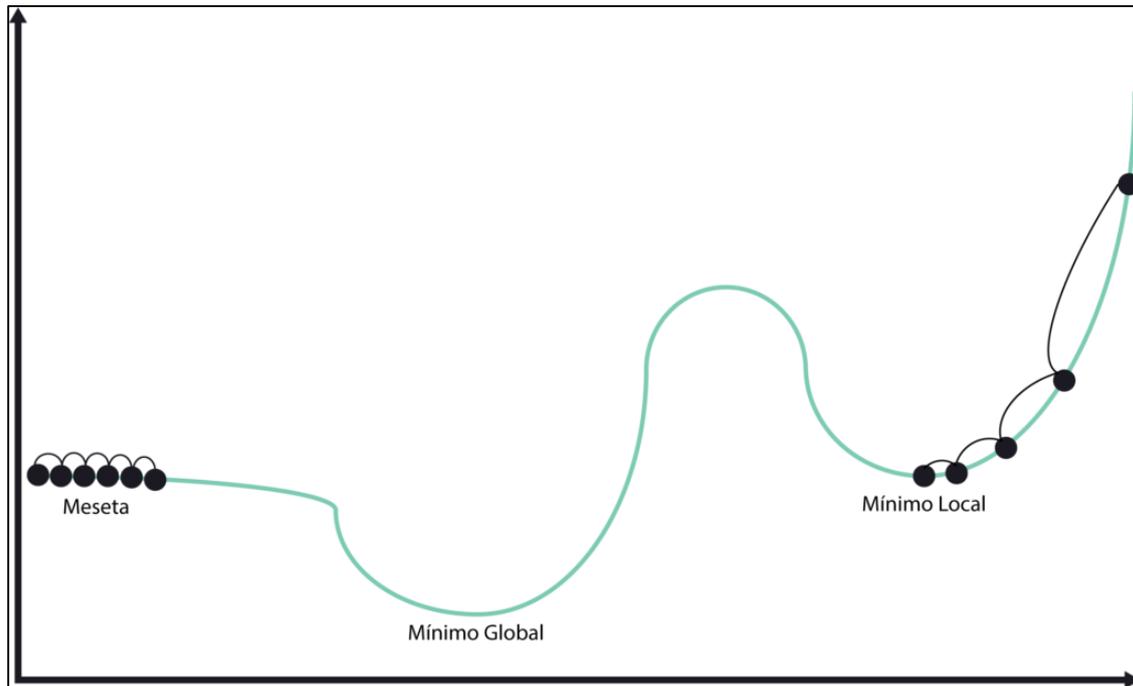
- Para minimizar la función de error, es crucial contrarrestar la gradiente, ya que esta indica la dirección de crecimiento. En esencia, nos desplazamos en la superficie del error con la intención de alcanzar su mínimo global. Sin embargo, existe el riesgo de quedar atrapados en mínimos locales si no se maneja adecuadamente esta técnica del gradiente descendente. Por esta razón, es prudente avanzar por la superficie de error con incrementos pequeños en los pesos, ya que, al tener información local, no se puede determinar la proximidad al punto mínimo. Con incrementos grandes, se corre el peligro de pasar por alto el punto mínimo, mientras que, con incrementos pequeños, aunque se tarde más en llegar, se previene esta eventualidad.

El incremento seleccionado juega un papel crucial en la velocidad de convergencia del algoritmo, y esta velocidad está bajo la influencia del factor de aprendizaje γ . Comúnmente, γ se elige como un número pequeño para asegurar que la red encuentre una solución adecuada. Un valor reducido de γ implica que la red requerirá un mayor número de iteraciones, mientras que un valor excesivamente grande generará cambios pronunciados en los pesos, moviéndose rápidamente en la superficie del error. Esto conlleva el riesgo de saltar el valor mínimo del error y quedar atrapado en oscilaciones alrededor de él, sin alcanzarlo.

Es aconsejable ajustar el valor de γ conforme disminuye el error durante la fase de entrenamiento. Esta estrategia garantiza una convergencia rápida, pero se debe tener precaución para evitar valores excesivamente grandes que puedan hacer que la red oscile y se aleje demasiado del valor mínimo.

Figura 10

Ilustración de Superficie de Error



Nota: Imagen extraída de (Rubio, 2024)

- b. Cuando se aplica el algoritmo de backpropagation para encontrar un mínimo, no existe garantía de que este sea un mínimo global. A pesar de ello, durante la estimación de la red, persistimos en utilizar el aprendizaje bajo este error, incluso cuando el valor es elevado.
- c. El algoritmo de backpropagation es sumamente sensible a la elección del factor de aprendizaje. Un valor muy pequeño (próximo a cero) resulta en variaciones diminutas en los pesos, garantizando significativamente la búsqueda. Por otro lado, si el factor es demasiado grande (cercano a 1), el valor de $R(q)$ puede disminuir con grandes oscilaciones, llevando a la red a saltar de un mínimo global a un mínimo local. Esto puede evitarse añadiendo un término $u\Delta w_{(r)}$

llamado momentum en la ecuación $\Delta w = -\gamma \frac{\partial R}{\partial w}$

$$\Delta w_{(r+1)} = -\gamma \frac{\partial R}{\partial w} + u\Delta w_{(r)} \quad (44)$$

La idea es dar a cada peso un impulso (inercia) de forma que la búsqueda del mínimo global continúe.

2.3.2. REDES NEURONALES RECURRENTE Y LSTM

Las Redes Neuronales Recurrentes (RNN) son una categoría de redes neuronales artificiales en las que las conexiones entre nodos forman un gráfico dirigido a lo largo de una secuencia temporal, permitiéndoles exhibir un comportamiento dinámico en el tiempo.

Estas redes pueden utilizar su estado interno o memoria para procesar secuencias de entradas de longitud variable, lo que las hace adecuadas para tareas como el reconocimiento de escritura conectada no segmentada o el reconocimiento de voz (Sezer, 2020).

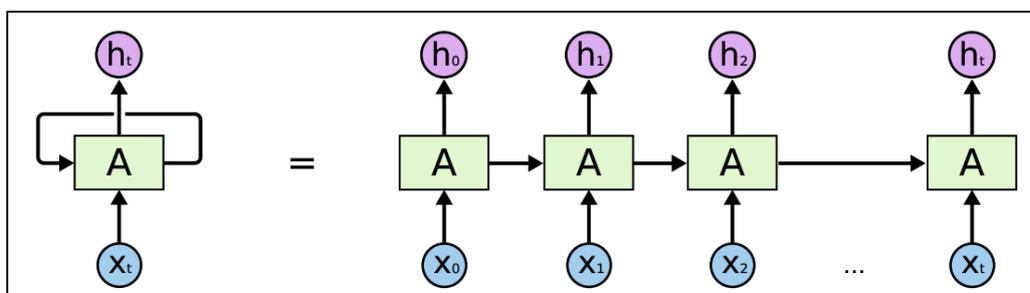
Las RNN se dividen en dos clases principales con una estructura general similar: las de impulso finito y las de impulso infinito. Es importante destacar que ambas clases muestran un comportamiento dinámico temporal (Kelleher, 2019).

Una red recurrente de impulso finito es un gráfico acíclico dirigido que puede desenrollarse y reemplazarse con una red neuronal estrictamente feedforward. Por otro lado, una red recurrente de impulso infinito es un gráfico cíclico dirigido que no puede desenrollarse (Sezer, 2020).

Tanto las redes recurrentes de impulso finito como las de impulso infinito pueden incorporar estados almacenados adicionales, los cuales pueden ser controlados directamente por la red neuronal. Este almacenamiento también puede ser reemplazado por otra red que incluya retardos temporales o bucles de retroalimentación. A estos estados controlados se les conoce como estado de compuerta o memoria de compuerta, también denominado Red Neuronal de Retroalimentación (Sezer, 2020).

Figura 11

Ejemplo de una red neuronal recurrente



Nota. Ilustración extraída del libro *Deep Learning with Keras: Implementing deep learning models and neural networks with the power of Python* (Gulli & Pal, 2017)

Las redes neuronales completamente recurrentes presentan una topología en la que las salidas de todas las neuronas se conectan con las entradas de todas las demás neuronas. Esta estructura es la más general en las redes neuronales y permite representar todas las demás topologías al establecer algunos pesos de conexión en cero para simular la falta de conexiones entre ciertas neuronas (Goodfellow, 2016).

En las topologías prácticas, las redes neuronales a menudo se organizan en capas, y al tener diferentes pasos en el tiempo, se vuelven completamente recurrentes.

La arquitectura de Long Short-Term Memory (LSTM), que aborda la memoria a largo y corto plazo, es utilizada en el aprendizaje profundo de modelos secuenciales con variables continuas, especialmente eficaz para la sincronización (Hua, 2019). A diferencia de las redes neuronales de retroalimentación estándar, las LSTM incorporan una conexión de retroalimentación y pueden procesar no solo un único punto de datos, sino todo el flujo de datos (Kelleher, 2019).

Las LSTM encuentran aplicaciones en diversas tareas, como el reconocimiento de escritura a mano sin particiones vinculadas, el reconocimiento de voz, la detección de anomalías en el tráfico de red o el sistema de detección de intrusiones, la clasificación de spam, el reconocimiento de géneros musicales y el análisis de sentimientos. Además, se han utilizado en el procesamiento del lenguaje natural para modelar árboles de dependencia, expresiones de oraciones, modelado del lenguaje y respuestas a preguntas (Sezer, 2020).

Una unidad LSTM típica consta de celdas, entradas, salidas y puertas de olvido. La celda almacena valores para períodos de tiempo arbitrarios, y las tres puertas regulan el flujo de información dentro y fuera de la celda. Las LSTM representan una extensión del modelo RNN clásico, siendo consideradas más dinámicas y exitosas en aplicaciones comerciales y financieras (Sezer, 2020).

La esencia clave de LSTM radica en la extensión del modelo RNN tradicional mediante la incorporación de bloques de memoria adicionales o estados de celda. El objetivo central es que el modelo aprenda a mantener su "memoria celular" a lo largo del tiempo, posibilitando la actualización constante de su estado interno. Para lograr esto, se introducen una o más puertas que permiten la adición o eliminación de información en la "memoria de la celda" en cada paso temporal (Hua, 2019).

A pesar de que los modelos basados en LSTM se consideran actualizados en el ámbito del aprendizaje automático supervisado, las redes neuronales recurrentes (RNN) en general presentan una dependencia persistente entre los datos en el

modelado. Las LSTM son especialmente útiles para realizar predicciones basadas en datos de series temporales, ya que pueden manejar uno o más retrasos desconocidos entre eventos significativos en dichas series (Tablada, sf). El desarrollo de las LSTM surge como respuesta a los problemas de "gradación explosiva" y desvanecimiento que pueden surgir durante el entrenamiento tradicional de RNN, derivados de una alta dependencia entre las neuronas, dando lugar a la pérdida de información (Hua, 2019). Sean matrices W_q y U_q que contienen, respectivamente, los pesos de las conexiones de entrada y se tomará que $c_t \in \mathbb{R}^h$ será una celda que contiene h celdas de unidad (Hua, 2019)

Para el caso de LSTM con una puerta para olvidar las ecuaciones para el pase directo de una unidad LSTM con una puerta de olvido los valores iniciales son $c_0 = 0$ y $h_0 = 0$. Teniendo las siguientes variables:

$x_t \in \mathbb{R}^d$: vector de entrada a la unidad

$L f_t \in \mathbb{R}^h$: que será la orden de olvidar el vector de activación de la puerta

$i_t \in \mathbb{R}^h$: entrada/ actualización del vector de activación de la puerta

\mathbb{R}^h : vector de activación de la puerta de salida

$h_t \in \mathbb{R}^h$: vector de estado también conocido como vector de salida de la unidad LSTM

$\tilde{c}_t \in \mathbb{R}^h$: vector de entrada de celda

\mathbb{R}^h : vector de entrada de celda

$W \in \mathbb{R}^{h \times d}$, $U \in \mathbb{R}^{h \times h}$ y $b \in \mathbb{R}^h$: matrices de peso y parámetros de vector de sesgo que deben aprenderse durante el entrenamiento.

Donde los superíndices d y h se refieren al número de funciones de entrada y al número de unidades ocultas, respectivamente.

Donde se pueden denotar las siguientes funciones de activación a escoger:

σ_g : función sigmoidea

σ_c : función tangente hiperbólica

σ_h : función de tangente hiperbólica o, LSTM de mirilla $\sigma_h(x) = x$

Ahora bien, las conexiones de mirilla LSTM

Las conexiones de mirilla permiten que las puertas accedan al "carrusel" de error constante (CEC), cuya activación es el estado de la celda $h_{\{t-1\}}$ usando lo siguiente:

$$f_t = \sigma_g(W_f x_t + U_f c_{t-1} + b_f)$$

$$i_t = \sigma_g(W_i x_t + U_i c_{t-1} + b_i)$$

$$o_t = \sigma_g(W_o x_t + U_o c_{t-1} + b_o)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \sigma_e(W_c x_t + b_c)$$

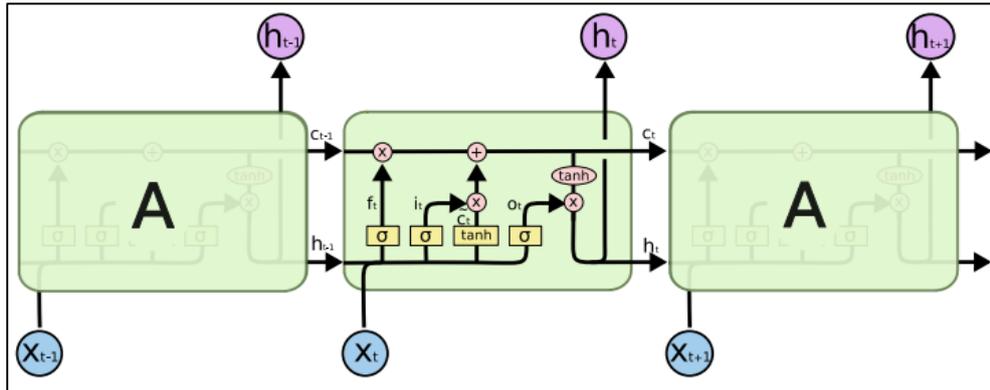
$$h_t = o_t \circ \sigma_h(c_t)$$

Ahora bien, también pueden tener una "mirilla convolucional" LSTM. El * denota el operador de convolución. Por lo que se tendría:

$$f_t = \sigma_g(W_f * x_t + U_f * h_{t-1} + V_f \circ c_{t-1} + b_f)$$

Figura 12

Ejemplo de una celda LSTM



Nota. Ilustración extraída del libro *Deep Learning with Keras: Implementing deep learning models and neural networks with the power of Python* (Gulli & Pal, 2017)

$$f_t = \sigma_g(W_f * x_t + U_f * h_{t-1} + V_f \circ c_{t-1} + b_f)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c * x_t + U_c * h_{t-1} + b_c)$$

$$o_t = \sigma_g(W_o * x_t + U_o * h_{t-1} + V_o \circ c_t + b_o)$$

$$h_t = o_t \circ \sigma_h(c_t)$$

resumiendo, a su forma simple puede ser presentada de la siguiente manera:

Los bloques de la célula de la red neuronal LSTM. (Hua, 2019) En el momento t , la unidad de memoria interna anota toda la información histórica hasta el momento actual y está controlada por tres "puertas o accesos":

1. La entrada a la unidad de memoria, representada como g_i^t (para el tiempo t y la celda i_1 , y de manera similar para otros símbolos), está regulada por la puerta de entrada. Esta puerta controla la incorporación de nueva información en la memoria interna:

$$g_i^t = \sigma \left(b_i^q + \sum_j U_{i,j}^q x_j^t + \sum_l W_{i,q}^q h_j^{t-1} \right)$$

Donde la función σ es la función sigmoide, $x^{(t)}$ es el vector de entrada actual $h^{(t)}$ es el vector de capa actual, que contiene la salida de todas las celdas LSTM y^g, U^g y W^g indican el sesgo, los pesos de entrada y los pesos recurrentes para la puerta de entrada, respectivamente.

2. Puerta de olvido: donde se olvida de la unidad de puerta $f_i^{(t)}$ gestiona la unidad de memoria interna sobre cuanta información en los momentos anteriores necesita ser guardada (Goodfellow, 2016)

$$f_i^t = \sigma \left(b_i^f + \sum_j U_{i,j}^f x_j^{(t)} + \sum_l W_{i,q}^f h_j^{(t-1)} \right)$$

donde la función de activación σ es la función sigmoidea con un valor entre 0 y 1 b^f, U^f, W^f son el sesgo, las ponderaciones de entrada y las ponderaciones recurrentes para la puerta de olvido, respectivamente.

Cuando $f_i^{(t)} = 1$; se abre la puerta de olvido y el estado de la celda en el momento se introduce en la celda.

De lo contrario, cuando $f_i^{(t)} = 0$; se cierra la puerta de olvido y se desprecia el estado de la celda en el momento inmediatamente anterior.

La unidad de estado interno $S_i^{(t)}$ se actualiza, que se formule como:

$$S_i^{(t)} = f_i^{(t-1)} S_i^{(t-1)} + g_i^{(t)} \sigma \left(b_i + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} h_j^{(t-1)} \right)$$

Donde b_i, U_i, W_i representan sesgo, ponderaciones de entrada y ponderaciones recurrentes en la celda LSMT, respectivamente.

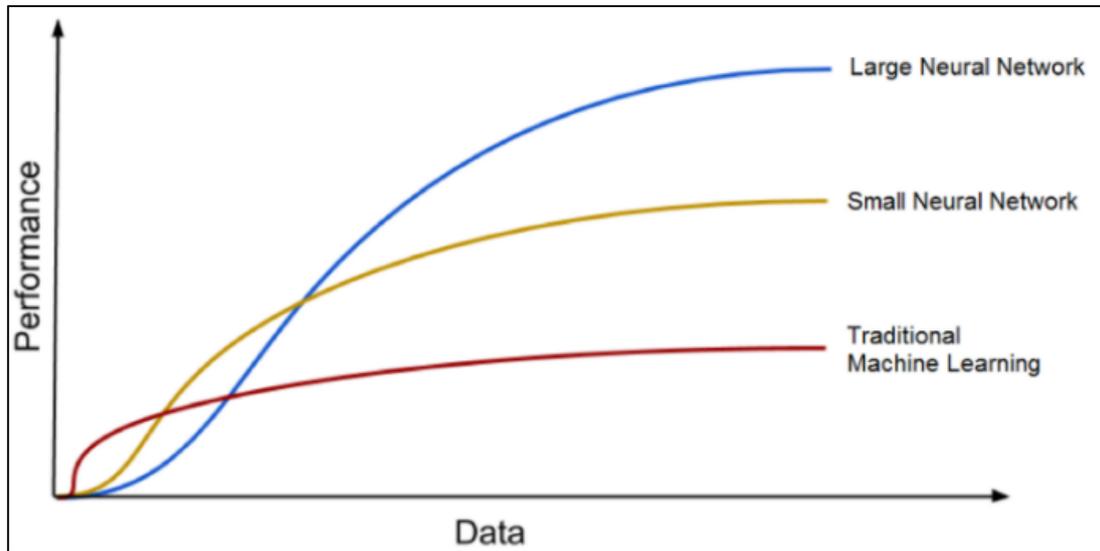
Denotando que la primera parte es la información del estado de la celda controlada por la puerta de olvido en el último momento, y la última parte es la información de entrada controlada por la puerta de entrada (Goodfellow, 2016).

3. Puerta de salida: La puerta de salida $O_i^{(t)}$ controla la unidad de memoria interna, es decir produce y reproduce la información requerida, que puede ser dada por:

$$h_i^{(t)} = \tanh(S_i^{(t)}) O_i^{(t)}$$

Figura 13

Rendimiento de modelos



Nota. Ilustración extraída del libro *Deep Learning with Keras: Implementing deep learning models and neural networks with the power of Python* (Gulli & Pal, 2017)

$$o_i^{(t)} = \sigma \left(b_i^0 + \sum_j U_{i,j}^0 h_j^{(t-1)} \right)$$

Donde b^0 , U^0 , W^0 denotan sesgo, pesos de entrada y pesos recurrentes para la puerta salida, respectivamente.

La salida de la unidad LSTM funciona como el estado de la capa oculta en la red neuronal recurrente (RNN). No hay una regla general para determinar el número óptimo de capas ocultas y retrasos, lo cual desempeña un papel crucial en la influencia del rendimiento durante el entrenamiento y las pruebas en la arquitectura LSTM. La elección de la arquitectura adecuada puede ser explorada mediante la prueba de varias configuraciones con un conjunto de datos específico para identificar aquellas que ofrecen un rendimiento óptimo (Goodfellow, 2016).

Mientras que los modelos tradicionales de machine learning y series temporales tienden a destacarse en la resolución de problemas con conjuntos de datos pequeños, las redes neuronales demuestran un mejor rendimiento en contextos con grandes cantidades de datos. La eficiencia de la red neuronal también se incrementa al aumentar la complejidad de la arquitectura y al incorporar conjuntos de datos más extensos (Goodfellow, 2016).

2.4. SERIES TEMPORALES

Un proceso estocástico es una colección de variables aleatorias indexadas por un parámetro, las cuales están generalmente interrelacionadas. Cuando dicho parámetro corresponde al tiempo y las variables aleatorias están secuenciadas cronológicamente, el proceso se denomina serie temporal (Anderson, 1985). De manera más precisa, una serie temporal es un conjunto de variables aleatorias dispuestas en orden temporal, es decir,

$$\{X_t\}_{-\infty}^{\infty} = \{\dots, X_{-2}, X_{-1}, X_0, X_1, X_2, \dots, X_s, X_{s+1}, \dots\}$$

Una realización de un proceso estocástico, y específicamente de una serie temporal, es un conjunto de valores u observaciones en el que, para cada instante t , la variable Y_t del proceso asume uno de estos valores. Así, una realización de $\{Y_t\}$ se configura como una secuencia de números reales o, de manera equivalente,

$$\{\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots, x_s, x_{s+1}, \dots\}$$

Es crucial destacar que el término "series de tiempo" suele emplearse tanto para referirse a la realización como al proceso estocástico subyacente que la genera. En numerosas aplicaciones prácticas, particularmente en el ámbito de las variables económicas, es común que no se disponga de múltiples realizaciones de una misma serie temporal; generalmente, se observa solo una única realización. Esta realización está compuesta por un número finito de observaciones, lo que implica que solo se tiene acceso a un subconjunto finito de las variables aleatorias del proceso estocástico en cuestión (Anderson, 1985). Por lo tanto, cuando se menciona una serie de tiempo, se alude comúnmente a los valores que toma un subconjunto finito de variables aleatorias, es decir,

$$\{y_1, y_2, \dots, y_T\}$$

2.4.1. COMPONENTES DE UNA SERIE TEMPORAL

El análisis clásico de las series temporales asume que los valores observados resultan de la interacción de tres componentes principales (Uriel Jimenez, 2012):

- a. Componente de tendencia: Este se define como la variación a largo plazo en el nivel medio de la serie, reflejando un cambio sostenido en la media a lo largo del tiempo. La tendencia se caracteriza por un movimiento suave y prolongado en la serie temporal (Uriel Jimenez, 2012).
- b. Componente estacional: Muchas series temporales exhiben periodicidad o variaciones cíclicas, como fluctuaciones semestrales o mensuales. Un

ejemplo es el aumento en las Ventas al Detalle en Puerto Rico durante noviembre y diciembre debido a las festividades navideñas. Estos efectos estacionales son comprensibles y medibles, y pueden ser eliminados del conjunto de datos mediante un proceso conocido como desestacionalización de la serie (Uriel Jimenez, 2012).

- c. Componente aleatoria: Esta componente no sigue ningún patrón específico y resulta de factores fortuitos o aleatorios que afectan la serie temporal de manera aislada.

De los tres componentes mencionados, los dos primeros son determinísticos, mientras que el último es aleatorio. Así, una serie temporal puede representarse como la suma de estas componentes (Peña, 2000).

$$X_t = T_t + E_t + I_t$$

Donde T_t es la tendencia, E_t es la componente estacional e I_t es la componente aleatoria.

2.4.2. CLASIFICACIÓN DESCRIPTIVA DE LAS SERIES TEMPORALES

Las series temporales se pueden clasificar en las siguientes categorías:

- a. Estacionarias: Una serie temporal es estacionaria si sus propiedades estadísticas, como la media y la varianza, son constantes a lo largo del tiempo. Gráficamente, esto se manifiesta en que los valores de la serie oscilan alrededor de una media constante, con una variabilidad que también permanece constante (Peña, 2000).
- b. No estacionarias: Estas series presentan cambios en la tendencia y/o la variabilidad a lo largo del tiempo. Las variaciones en la media indican una tendencia de crecimiento o decrecimiento a largo plazo, lo que implica que la serie no oscila alrededor de un valor constante (Peña, 2000).

2.4.3. PROCESO ESTOCÁSTICO ESTACIONARIO

Un proceso estocástico se considera estacionario si su media y varianza permanecen constantes a lo largo del tiempo, y si la covarianza entre dos periodos depende únicamente de la distancia o el rezago entre estos periodos, y no del momento específico en que se calcula la covarianza (Peña, 2000).

Sea X_t una serie de tiempo entonces con estas propiedades:

$$\text{Media } E(X_t) = E(X_{t+k}) = \mu$$

$$\text{Varianza } V(X_t) = V(X_{t+k}) = \sigma^2$$

$$\text{Covarianza } \gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)]$$

donde γ_k , la covarianza al rezago k, es la covarianza entre los valores de X_t y X_{t+k} , que están separados k periodos (Peña, 2000).

2.4.4. RUIDO BLANCO

El ruido blanco es un caso elemental de proceso estocástico, caracterizado por valores que son independientes e idénticamente distribuidos en el tiempo, con una media de cero y varianza constante (Peña, 2000), se denota por ε_t .

$$\varepsilon_t \sim N(0, \sigma^2) \quad \text{cov}(\varepsilon_{t_i}, \varepsilon_{t_j}) = 0 \quad \forall t_i \neq t_j$$

2.4.5. FUNCIÓN DE AUTOCORRELACIÓN

Debido a que la covarianza puede presentar problemas relacionados con las unidades de medida de las variables aleatorias, es común emplear el coeficiente de correlación en su lugar. Este coeficiente se obtiene al estandarizar la covarianza, eliminando así las unidades y facilitando la comparación (Peña, 2000). La correlación de orden k, ρ_k , $k = \dots, -2, -1, 0, 1, 2, \dots$, es definida como

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t+k})}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{var}(Y_{t+k})}}$$

Cuando el proceso es estacionario en covarianza

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Tanto la covarianza como la correlación evalúan la relación entre las variables aleatorias dentro del mismo proceso, específicamente, la relación de una variable en diferentes instantes de tiempo (Peña, 2000). Por lo tanto, se les denomina auto covarianza y autocorrelación, respectivamente.

2.4.6. MODELOS DETERMINÍSTICOS

Los modelos de tendencia determinista, una ampliación directa de los métodos de regresión, son útiles para describir los patrones en series temporales, pero sus predicciones suelen ser imprecisas, con un alto margen de error. Esto se debe a que, en series temporales, las observaciones recientes están más influenciadas por sus valores pasados, con una dependencia decreciente hacia datos más antiguos. Los

modelos de tendencia determinista no aprovechan esta propiedad (Bowerman & Koehler, 2007).

Dentro de las técnicas univariantes, hay métodos simples como modelos autorregresivos de primer orden o modelos de tendencia lineal o exponencial. Sin embargo, hay técnicas más complejas como los modelos Box-Jenkins o los modelos de función de transferencia. “En teoría, la metodología de Box-Jenkins se puede usar para pronosticar cualquier tipo de serie de tiempo. Es la más adecuada cuando los componentes de la serie de tiempo están cambiando conforme transcurre el tiempo” (Bowerman & Koehler, 2007)

2.5. METODOLOGÍA BOX JENKINS

El método predictivo Box-Jenkins es una estrategia iterativa que consta de cuatro fases fundamentales:

1. Identificación del modelo: Se inicia evaluando la estacionariedad de la serie, utilizando un gráfico de dispersión. Si la serie no es estacionaria, se recurre a la diferenciación para lograr estacionarizarla. Box-Jenkins convierte los datos en una serie estacionaria y realiza los cálculos subsiguientes con los datos transformados (Uriel Jimenez, 2012)
2. Estimación y prueba de adecuación: Una vez identificados los valores apropiados de p , d y q , la siguiente fase implica la estimación de los parámetros de los términos autorregresivos y de media móvil presentes en el modelo. La estimación puede realizarse mediante mínimos cuadrados simples o, en algunos casos, se puede optar por métodos de estimación no lineal de parámetros. Diversos paquetes estadísticos facilitan este cálculo, lo que minimiza la necesidad de involucrarse en desarrollos matemáticos complejos (Ramos, 2017).
3. Verificación de diagnóstico: Una vez seleccionado un modelo ARIMA específico y estimados sus parámetros, se procede a verificar si dicho modelo se ajusta de manera razonable a los datos. Dada la posibilidad de que existan otros modelos ARIMA adecuados, la elección del modelo se considera más un arte que una ciencia, requiriendo habilidad para la selección correcta (Uriel Jimenez, 2012)

4. Predicción: La última fase destaca por la eficacia de ARIMA en la predicción, siendo especialmente exitoso para pronósticos a corto plazo. Las predicciones obtenidas mediante este método suelen ser más confiables que las obtenidas mediante enfoques de modelado tradicionales. Una vez definido el modelo, se procede a realizar pronósticos futuros utilizando el modelo obtenido en el procedimiento (Ramos, 2017).

2.5.1. MODELOS DE PROMEDIO MÓVIL AUTORREGRESIVO INTEGRADO: ARIMA (P,D,Q).

Se trata de un enfoque para modelar que se basa en la descripción de un valor como una función lineal de datos previos y errores aleatorios, siendo aplicado específicamente a series temporales estacionarias. Los modelos ARIMA (Autorregresivo Integrado de Promedio Móvil) constituyen una categoría especializada de técnicas de filtración que prescinden completamente de las variables independientes en la formulación de pronósticos. Estos modelos, altamente sofisticados en cuanto a ajuste de curvas, aprovechan tanto los valores anteriores como reales de la variable dependiente para generar pronósticos precisos a corto plazo (Ramos, 2017)

En 1970, Box y Jenkins crearon un marco metodológico con el propósito de reconocer, estimar y evaluar modelos dinámicos de series temporales, donde la variable temporal desempeña un papel esencial (Fernandez, 2021). La atención centrada únicamente en los valores previos de una variable específica para explicar su evolución actual y futura se presenta simultáneamente como una ventaja y un desafío:

- La ventaja reside en prescindir de la necesidad de múltiples conjuntos de datos (variables distintas) referidos al mismo periodo de tiempo, una característica común en todos los modelos univariantes. Al mismo tiempo, se evita la identificación y especificación del modelo en el sentido de la econometría tradicional (Anderson, 1985).
- El inconveniente radica en que, al omitir la inclusión de un conjunto más amplio de variables explicativas, no se consideran las relaciones que seguramente existen entre casi todas las variables económicas. Esto resulta en una pérdida de capacidad de análisis temporal, ya que implícitamente renunciamos al estudio teórico previo del fenómeno y a su innegable utilidad (Anderson, 1985).

Los modelos $ARIMA(p, d, q)$ constituyen una clase particular de procesos no estacionarios y se define como:

$$W_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} - \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \dots - \omega_q \varepsilon_{t-q}$$

donde:

$W_t = Y_t - Y_{t-1}$: primera diferencia

d : Número de diferenciaciones.

El análisis de series temporales, basado en el modelo ARMA, se enfrenta al desafío de tratar con procesos no estacionarios, una característica común en la mayoría de las series estadísticas. Para abordar este problema, es crucial explorar mecanismos de transformación que conviertan estas series en procesos estacionarios (Ramos, 2017). Aunque existen diversas formas en las que la no estacionariedad puede introducirse en un proceso estocástico (potencialmente infinitas), nos enfocaremos en aquellas que son pertinentes para describir el comportamiento de las series económicas y, al mismo tiempo, susceptibles de transformarse en procesos estacionarios (Ramos, 2017). Comenzaremos examinando el concepto de "caminata aleatoria".

Caminata aleatoria

Ramos, en (2017) definió como Y_t es una serie de tiempo estocástica caracterizada por cambios sucesivos, representados como u_t , que se extraen de manera independiente de una distribución de probabilidad con una media de 0 y una varianza de σ^2 . En consecuencia, la determinación de Y_t sigue la siguiente relación:

$$Y_t = Y_{t-1} + \varepsilon_t$$

donde ε_t es un ruido blanco.

Caso general

Dada una serie Y_t , que eventualmente corresponde a los logaritmos de los valores originales, si su diferencia de orden "d" puede ser representada por un proceso $ARMA(p, q)$ estacionario, se dice que la serie Y_t sigue un proceso $ARIMA(p, d, q)$. La letra "I" en ARIMA corresponde a "Integración", la operación inversa a la diferenciación (Anderson, 1985)

Si $Z_t = \Delta^d Y_t$ y Z_t sigue un proceso $ARMA(p, q)$ estacionarios:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) Z_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t$$

$$\phi_p(L) Z_t = \theta_q(L) \varepsilon_t$$

entonces Y_t sigue un proceso $ARIMA(p, d, q)$. También se escribe en la variable original Y_t como:

$$\phi_p(L)(1-L)^d Y_t = \theta_q(L)\varepsilon_t$$

generalmente, no son necesarias diferencias regulares de orden superior a 2, excepto en el caso de variables que presentan estacionalidad (Peña, 2000)

2.5.2. MODELOS ARIMA ESTACIONALES

Un modelo estacional puro se distingue por establecer conexiones solo entre observaciones que tienen una separación de s periodos o múltiplos de s (Uriel Jimenez, 2012). Estas series exhiben ciclos u oscilaciones estrictamente periódicas, donde el periodo es igual o inferior a un año. La notación para el periodo estacional es "s", siendo $s = 4$ para datos trimestrales y $s = 12$ para datos anuales. El desarrollo de modelos ARIMA estacionales comparte características similares con los modelos ARIMA no estacionales. Según Anderson (1985), los métodos que emplean modelos estacionales o $SARIMA(p, d, q)(P, D, Q)$ presuponen que el componente estacional es generado por un proceso estocástico. La identificación de este proceso se lleva a cabo de manera similar a los modelos que representan la estructura regular de una serie, con la diferencia de que se examinan los "valores estacionales" de las funciones de autocorrelación (correspondientes a rezagos 4, 8, 12, ... en datos trimestrales y 12, 24, 36, ... en datos mensuales) (Anderson, 1985). De esta manera, una serie podría requerir diferencias de orden si los valores estacionales de la función de autocorrelación no tienden a cero rápidamente. Un proceso $SARIMA(p, d, q)(P, D, Q)$ se define así,

$$\phi_p(L)\Phi_P(L^s)(1-L)^d(1-L^s)^D Y_t = \theta_q(L)\Theta_Q(L^s)\varepsilon_t$$

donde

$$\phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

$$\theta_q(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$$

$$\Phi_P(L^s) = 1 - \Phi_1 L^s - \Phi_2 L^{2s} - \dots - \Phi_P L^{Ps}$$

$$\Theta_Q(L^s) = 1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_Q L^{Qs}$$

2.5.3. PRUEBAS PARA ANALIZAR LA TENDENCIA

2.5.3.1. Prueba de Mann Kendall (MK)

La prueba de Mann-Kendall (MK) es una prueba estadística no paramétrica que se fundamenta en el análisis de rangos para determinar la significancia de una tendencia. Ha sido ampliamente aplicada en la detección de tendencias en investigaciones hidrológicas. La hipótesis nula H_0 : No existe una tendencia en la serie $(X_i, i = 1, 2, 3 \dots n)$. La hipótesis alternativa es H_a : Hay una tendencia en la serie (Crespo, 2016).

Su desarrollo se resume como:

1. Se listan los valores de las variables, de forma ordenada (x_1, x_2, \dots, x_n)
2. Se calcula el signo de la diferencia entre cada par de valores al comparar sus magnitudes. $(x_j - x_k)$ con $(j > k)$ de acuerdo con lo siguiente:

$$\text{signo}(x_j - x_k) = \begin{cases} 1 & \text{si } (x_j - x_k) > 0 \\ 0 & \text{si } (x_j - x_k) = 0 \\ -1 & \text{si } (x_j - x_k) < 0 \end{cases}$$

3. Obtención del estadístico S de Mann Kendall, mediante la ecuación

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{signo}(x_j - x_k)$$

Si el valor resultante, denotado como S, es positivo, se deduce de manera subjetiva que existe una tendencia creciente, mientras que si S es negativo, se infiere que la tendencia es decreciente.

4. Con base a los indicadores se estima una varianza para el estadístico S de Mann Kendall, que considera el caso de los empates $(x_j - x_k = 0)$ obtenidos en el paso 2, mediante la ecuación:

$$\text{Var}(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{q=1}^g t_q(t_q-1)(2t_q+5) \right]$$

5. Cálculo del estadístico Z_{MK} mediante:

$$Z_{MK} = \begin{cases} \frac{S-1}{[\text{Var}(S)]^{1/2}} & \text{si } > 0 \\ 0 & \text{si } = 0 \\ \frac{S+1}{[\text{Var}(S)]^{1/2}} & \text{si } < 0 \end{cases}$$

A partir del estadístico Z_{MK} se evalúa la hipótesis de interés, que puede ser:

a) H_0 : No hay tendencia vs. H_1 : Hay tendencia decreciente

b) H_0 : No hay tendencia vs. H_1 : Hay tendencia creciente

La prueba de Mann-Kendall nos ofrece un valor de significancia denominado p-value. Si el valor de p es menor que 0.05, podemos afirmar con alta confianza que la hipótesis alternativa es verdadera.

2.5.3.2. Prueba de Cox y Stuart

La evaluación de la significancia de la tendencia en los datos se lleva a cabo principalmente mediante la prueba de Mann-Kendall, tal como se describe en el marco teórico. Dependiendo de la presencia de autocorrelación en los datos, se elige entre la versión modificada o la prueba estándar. Cuando no hay autocorrelación, se realiza la prueba de tendencia Cox and Stuart como procedimiento de confirmación, basada en un principio similar a la prueba de Mann-Kendall. Posteriormente, se llega a una conclusión sobre la significancia de la tendencia en la serie (Martínez, 2004).

Asimismo, al igual que la prueba del signo utiliza la distribución binomial para evaluar la hipótesis de una distribución idéntica versus diferencias en la posición, es posible crear variables binomiales adicionales para evaluar la hipótesis de aleatoriedad versus alternativas de tendencia. Esta prueba es una alternativa a la prueba paramétrico para $H_0: \beta = 0$ en el modelo de regresión lineal $Y = \alpha + \beta X + \varepsilon$. La hipótesis nula en esta prueba implica que la pendiente de la recta es 0 (Martínez, 2004).

La prueba de Cox-Stuart, fundamentada en variables aleatorias binomiales, se emplea para evaluar la presencia de tendencias. A diferencia de las pruebas paramétricas clásicas, que contrastan la ausencia de tendencia con la presencia de una tendencia lineal (una relación lineal entre X e Y), la prueba de Cox-Stuart no impone esa restricción y contrasta la hipótesis de ausencia de tendencia con la hipótesis alternativa de tendencia monótona. Una tendencia es considerada monótona si la variable dependiente aumenta cuando la variable independiente aumenta (monótona creciente) o disminuye cuando la variable independiente aumenta (monótona decreciente) (Martínez, 2004).

Sea x_1, \dots, x_n una sucesión de v.a. ordenadas según algún criterio. Si los x_i están aleatoriamente distribuidos sobre la sucesión, entonces

$$P(x_i - x_j < 0) = P(x_i - x_j > 0) = \frac{1}{2} \quad \text{para todo } i \neq j$$

En cambio, si x_i crece monótonamente, $P(X_i - X_j < 0) > P(X_i - X_j > 0)$ si $i < j$ y si X_i decrece monótonamente, $P(X_i - X_j < 0) < P(X_i - X_j > 0)$ si $i < j$.

2.5.3.3. Correlación de Spearman

Tras revisar diversos autores, adoptamos el concepto de Spearman (Rho de Spearman). Este coeficiente representa una medida de asociación lineal que emplea los rangos, es decir, los números de orden, de cada grupo de sujetos y los compara. Existen dos métodos para calcular este coeficiente de correlación de rangos: uno propuesto por Spearman y otro por Kendall. El coeficiente Rho de Spearman, también conocido como R de Spearman, es más sencillo de calcular que el coeficiente de Kendall (Martínez, 2004).

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

en donde $d_i = r_{xi} - r_{yi}$ es la diferencia entre los rangos de X e Y .

2.5.3.4. Prueba de tendencias de Wald-Wolfowitz

La prueba de tendencias de Wald-Wolfowitz es un método no paramétrico que compara las distribuciones de dos poblaciones utilizando muestras aleatorias independientes. En esta prueba, la hipótesis nula postula que las dos muestras aleatorias provienen de poblaciones con distribuciones idénticas. A diferencia de la prueba U de Mann-Whitney, la prueba de Wald-Wolfowitz no se limita a detectar diferencias solo en la tendencia central; su hipótesis alternativa es mucho más amplia, sugiriendo que las distribuciones pueden diferir en aspectos como la tendencia central, la dispersión o la asimetría (Canavos, 1988).

Al igual que en la prueba de Mann-Whitney, las observaciones en las dos muestras aleatorias se combinan y ordenan según sus magnitudes. Sin embargo, en lugar de considerar los rangos, el procedimiento de Wald-Wolfowitz busca el número de tendencias en la secuencia ordenada (Canavos, 1988).

Definición: En la definición, se establece que una tendencia de longitud j se refiere a una secuencia de J observaciones pertenecientes al mismo grupo, que están rodeadas por observaciones de un grupo diferente, ya sea precedidas o seguidas por estas (Canavos, 1988).

Consideremos r como el número total de tendencias observadas en una secuencia ordenada de $n_1 + n_2$ observaciones, donde n_1 y n_2 son los respectivos tamaños de las muestras. Los valores posibles de r son $2, 3, \dots, n_1 + n_2$. Se puede demostrar que la función de probabilidad de R está definida por

$$p(r) = \begin{cases} \frac{2 \binom{n_1 - 1}{\frac{r}{2} - 1} \binom{n_2 - 1}{\frac{r}{2} - 1}}{\binom{n_1 + n_2}{n_1}} & r \text{ par} \\ \frac{\binom{n_1 - 1}{\frac{r}{2} - 1/2} \binom{n_2 - 1}{\frac{r}{2} - 3/2} + \binom{n_1 - 1}{\frac{r}{2} - 3/2} \binom{n_2 - 1}{\frac{r}{2} - 1/2}}{\binom{n_1 + n_2}{n_1}} & r \text{ impar} \end{cases}$$

La media y la varianza de R son

$$E(R) = \frac{2n_1n_2}{n_1 + n_2} + 1$$

$$Var(R) = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

Para probar H_0 con una probabilidad α , para el error tipo I, debe encontrarse un entero r_0 tal, que en la medida de lo posible

$$\sum_{r=2}^{r_0} p(r) = \alpha$$

Se rechaza la hipótesis nula cuando el número observado de tendencia es menor o igual a r_0 . Nótese que la región crítica es una región unilateral inferior dado que se rechaza H_0 cuando el número de tendencias es bastante pequeño (Canavos, 1988).

La distribución acumulativa de R se encuentra tabulada en forma extensa; pero si tanto n_1 como n_2 son mayores que 10, la distribución de R se encuentra, en forma adecuada, aproximada por una distribución normal con media y varianza $E(R)$ y $Var(R)$ respectivamente (Canavos, 1988).

2.5.4. ANÁLISIS DE LA ESTACIONALIDAD MEDIANTE GRÁFICOS.

2.5.4.1. Gráfico de *monthplot* y *box plot*

El *monthplot*, también conocido como gráfico de subseries estacionales, es una representación alternativa que resalta los patrones estacionales. En este gráfico, los datos de cada temporada se organizan en mini gráficos de tiempo separados, lo que permite visualizar tanto las diferencias estacionales entre grupos como los patrones dentro de cada grupo.

Por otro lado, el diagrama de cajas es eficaz para mostrar las diferencias estacionales entre grupos, pero no proporciona información sobre los patrones dentro de cada grupo. Sin embargo, para conjuntos de datos extensos, el diagrama de cajas suele ser más fácil de interpretar que el *monthplot*.

2.5.4.2. Uso de *correlograma*

El gráfico de autocorrelación (ACF) es una herramienta útil para detectar la presencia de estacionalidad en los datos. Calcula la diferencia entre un valor de la serie y otro rezagado en el tiempo, lo que proporciona puntos cercanos cuando no hay estacionalidad y puntos distantes cuando hay discrepancia. Estos últimos indican un nivel de estacionalidad en los datos (Bowerman & Koehler, 2007).

2.5.4.3. Prueba de *kruskal wallis*

La prueba de Kruskal-Wallis, desarrollada por William Kruskal y W. Allen Wallis, es un método no paramétrico utilizado en estadística para determinar si un conjunto de datos proviene de la misma población. Conceptualmente, es similar al ANOVA, pero en lugar de datos numéricos, se utilizan categorías. Esta prueba es una extensión de la prueba U de Mann-Whitney para tres o más grupos (Martínez, 2004).

Al ser no paramétrica, la prueba de Kruskal-Wallis no requiere que los datos sigan una distribución normal, a diferencia del ANOVA tradicional. Sin embargo, bajo la hipótesis nula, asume que los datos provienen de la misma distribución. Una violación común de este supuesto ocurre con datos heterocedásticos.

El estadístico utilizado en esta prueba es

$$K = (N - 1) \frac{\sum_{i=1}^g n_i (\bar{r}_i - \bar{r})^2}{\sum_{i=1}^g \sum_{j=1}^{n_i} (\bar{r}_{ij} - \bar{r})^2}$$

Finalmente, el pvalue (valor p) es aproximado por $P(\chi_{g-1}^2) \geq K$. Si algún n_i es pequeño la distribución de k puede ser distinta del chi cuadrado (Quispe, et.al, 2019).

2.5.5. EVALUACIÓN DE LA BONDAD DEL MODELO PREDICTIVO

Después de realizar una predicción en una serie temporal, es fundamental evaluar qué tan bien se adapta el modelo utilizado para realizar dicha predicción a los datos observados. Para este propósito, se emplean diversas métricas que permiten cuantificar el ajuste del modelo a los datos. Estas métricas se centran exclusivamente en evaluar el ajuste del modelo y no la calidad de la predicción en sí misma. En general, cuanto menor sea el valor de estas métricas, mejor será el ajuste del modelo a los datos de la serie temporal. A continuación, se detallarán estas métricas (Calvo, 2009).

2.5.5.1. Error Cuadrático Medio (MSE) y Raíz del Error Cuadrático Medio (RMSE)

El error cuadrático medio (MSE) se calcula al restar los valores reales de los estimados, elevando al cuadrado esta diferencia y obteniendo la media de estos valores. Al calcular la raíz cuadrada del MSE, obtenemos el error cuadrático medio raíz (RMSE), que conserva las ventajas del MSE pero es más representativo y fácil de interpretar directamente, ya que expresa el error en las mismas unidades que la variable predicha. El RMSE es una de las métricas más simples y comunes para la evaluación de modelos de regresión. Es útil cuando se desea evaluar valores atípicos que puedan ser de interés, ya que penaliza con mayor severidad los errores grandes en comparación con los errores pequeños. Sin embargo, su principal desventaja radica en que puede ser problemático cuando se tienen datos muy ruidosos (Calvo, 2009).

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

2.5.5.2. Error Absoluto Medio (MAE)

El error absoluto medio (MAE) se calcula promediando las diferencias absolutas entre los valores observados y las predicciones. Es una medida lineal en la que todas las diferencias individuales se ponderan por igual. Al igual que el RMSE, proporciona el error en las mismas unidades que la variable predicha. A diferencia del MSE, el MAE no es tan sensible a los valores atípicos y no penaliza los errores grandes con la misma severidad, lo que lo hace menos restrictivo. Es una solución simple para cuantificar la inexactitud del error promedio (Calvo, 2009).

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

2.5.5.3. Error Porcentual Absoluto Medio (MAPE)

El error porcentual absoluto medio (MAPE) es la media de las diferencias absolutas entre los valores predichos y los verdaderos, dividida por el valor verdadero, como se muestra en la Tabla 7. Esta métrica es comúnmente utilizada debido a su facilidad de interpretación y explicación. Por ejemplo, un MAPE del 11,5% indica que la diferencia promedio entre el valor real y el valor predicho es el 11,5% del valor real. Un menor valor de MAPE indica una mayor precisión en las predicciones del modelo. Por ejemplo, un modelo con un MAPE del 2% tiene una precisión superior a uno con un MAPE del 10%. El MAPE funciona mejor en conjuntos de datos sin valores extremos y sin ceros, ya que no es válido dividir entre cero (Calvo, 2009).

$$MAPE = \frac{1}{n} \sum_{t=1}^n 100 \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

2.6. HIPOTESIS

2.6.1. HIPOTESIS GENERAL.

Las redes neuronales recurrente multicapa, predicen mejor las series temporales de denuncias policiales en la Región del Cusco para el período de 2015 a 2022, que las series con el modelo ARIMA.

2.6.2. HIPOTESIS ESPECIFICAS

- a) Existe variabilidad significativa en la naturaleza y evolución temporal de las denuncias policiales en la región del Cusco entre 2015 y 2022.

- b) Existe un patrón estacional en las denuncias policiales en la región del Cusco durante el periodo 2015-2022.
- c) Existe un modelo ARIMA específico que se ajusta óptimamente a la serie temporal de denuncias policiales en la provincia del Cusco.
- d) Las redes neuronales LSTM en series temporales ofrecerá una predicción que se ajusta a la serie de datos de denuncias policiales en la provincia del Cusco

2.7. IDENTIFICACION DE VARIABLES E INDICADORES

Variable estudio: *Denuncias policiales*

2.8. OPERACIONALIZACIÓN DE VARIABLES

Tabla 2. Operacionalización de variable

Variable	Definición conceptual	Definición operacional	Indicadores
Denuncias	Es el acto procesal consistente en una declaración de conocimiento (verbal o escrita) emitido por una persona determinada, en virtud de la cual proporciona al titular del órgano competente la noticia de la existencia de un hecho que reviste los caracteres de delito. (Ccopa, 2019)	Acción de avisar a las autoridades (policía) sobre faltas o la comisión de un delito.	Denuncias de los delitos i. Delitos contra la vida, el cuerpo y la salud ii. Delitos contra el honor iii. Delitos contra la familia iv. Delitos contra la libertad v. Delitos contra el patrimonio vi. Delitos contra la confianza y la buena fe en los negocios vii. Delitos contra los derechos intelectuales viii. Delitos contra el patrimonio cultural ix. Delitos contra el orden económico x. Delitos contra el orden financiero y monetario xi. Delitos tributarios xii. Delitos contra la seguridad pública

			<p>xiii. Delitos ambientales</p> <p>xiv. Delitos contra la tranquilidad publica</p> <p>xv. Delitos contra la humanidad</p> <p>xvi. Delitos contra el estado y la defensa nacional</p> <p>xvii. Delitos contra los poderes del estado y el orden constitucional</p> <p>xviii. Delitos contra la voluntad popular</p> <p>xix. Delitos contra la administración publica</p> <p>xx. Delitos contra la fe publica</p>
--	--	--	--

CAPÍTULO III

METODOLOGÍA

3.1. ÁMBITO DE ESTUDIO: LOCALIZACIÓN POLÍTICA Y GEOGRÁFICA

La superficie del departamento de Cusco es de 71 987 km² (5,6 por ciento del territorio nacional), constituyéndose en la quinta más extensa del Perú.

Este departamento se localiza en la parte sur-oriental del territorio nacional y limita con los departamentos de Junín y Ucayali por el norte, Madre de Dios y Puno por el este, Arequipa por el sur-oeste y, Apurímac y Ayacucho por el oeste. Cusco, la ciudad capital, está ubicada a 3 399 m.s.n.m (BCRP, 2015)

El departamento de Cusco está dividido políticamente en 13 provincias y 114 distritos

Figura 14.
Mapa del departamento del Cusco



Nota. Tomada del Informe técnico: estadísticas de la criminalidad, seguridad ciudadana y violencia (INEI, 2021)

3.2. TIPO Y NIVEL DE INVESTIGACIÓN.

3.2.1. ENFOQUE DE INVESTIGACIÓN

El estudio según Sampieri (2018) es Cuantitativo, debido a que se realizara mediciones sobre la variable de interés.

3.2.2. TIPO DE INVESTIGACIÓN

La investigación es Aplicada, Según Carrasco (2018), los tipos de investigación se clasifican en básica y aplicada. La investigación aplicada tiene por objetivo aplicar los conocimientos en temas de la realidad, en este sentido el presente trabajo es aplicado.

3.2.3. ALCANCE O NIVEL DE INVESTIGACIÓN

Según profundidad de la investigación es descriptiva, debido a que solo contamos con una variable como son las denuncias policiales.

3.2.4. DISEÑO DE INVESTIGACIÓN

Aquí no interviene la participación del investigador por lo tanto el estudio es no experimental.

3.3. UNIDAD DE ANALISIS

Denuncias policiales

3.4. POBLACION DE ESTUDIO

Todas las denuncias policiales registradas en la región del Cusco

3.5. TAMAÑO DE MUESTRA

El estudio fue censal, debido a que se trabajara con la totalidad de la data.

3.6. TECNICAS DE RECOLECCION DE INFORMACIÓN

La técnica de recolección es documental, debido a que para el presente estudio, se solicitaron facilidades de información a la Policía Nacional del Perú (PNP). En respuesta a esta solicitud, la PNP proporcionó una base de datos completa y detallada que incluye todas las denuncias registradas en la región del Cusco durante el período comprendido entre 2015 y 2022. Esta base de datos contenía información relevante, como la cantidad de las denuncias por comisaria por mes y por año, tipo de delito, ubicación geográfica.

3.7. TECNICAS DE ANALISIS E INTERPRETACIÓN DE LA INFORMACION.

El análisis de los datos se llevó a cabo mediante un enfoque multifacético, abordando tanto la estructura temporal como las características estadísticas de las denuncias policiales en la región del Cusco durante el período 2015-2022.

Se inició con la visualización gráfica de la serie temporal de las denuncias, lo cual permitió identificar patrones preliminares y anomalías. A continuación, se analizó la tendencia de la serie utilizando varias pruebas estadísticas robustas, como las pruebas de Mann-Kendall y Cox-Stuart, que son métodos no paramétricos adecuados para detectar tendencias monotónicas en series de datos. Asimismo, se empleó la correlación de Spearman para evaluar la relación monotónica entre el tiempo y el número de denuncias, así como la prueba de Wald-Wolfowitz como una técnica que prueba que la serie no se comporta de manera aleatoria (tiene tendencia).

Luego para identificar y cuantificar la estacionalidad, se utilizaron diversos gráficos y pruebas estadísticas. Se emplearon gráficos de mes a mes (monthplot) y diagramas de caja (boxplot) para visualizar las variaciones mensuales y estacionales de los datos. Además, se construyó un correlograma para identificar patrones de autocorrelación a distintos retardos temporales. La prueba de Kruskal-Wallis se aplicó para detectar diferencias significativas en las distribuciones de los datos entre diferentes meses o estaciones, proporcionando una evaluación no paramétrica de la estacionalidad.

Para modelar y predecir la evolución futura de las denuncias, se aplicaron modelos ARIMA (AutoRegressive Integrated Moving Average), una técnica clásica y robusta en el análisis de series temporales. Estos modelos permitieron descomponer la serie en componentes de tendencia, estacionalidad y ruido, proporcionando una base para la previsión.

Como parte del análisis avanzado y de la predicción, se implementaron Redes Neuronales Recurrentes (LSTM). Estas redes son especialmente adecuadas para capturar patrones complejos y dependencias a largo plazo en series temporales, permitiendo una modelización más precisa de los datos y mejorando la capacidad de predicción en comparación con los métodos tradicionales como ARIMA.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

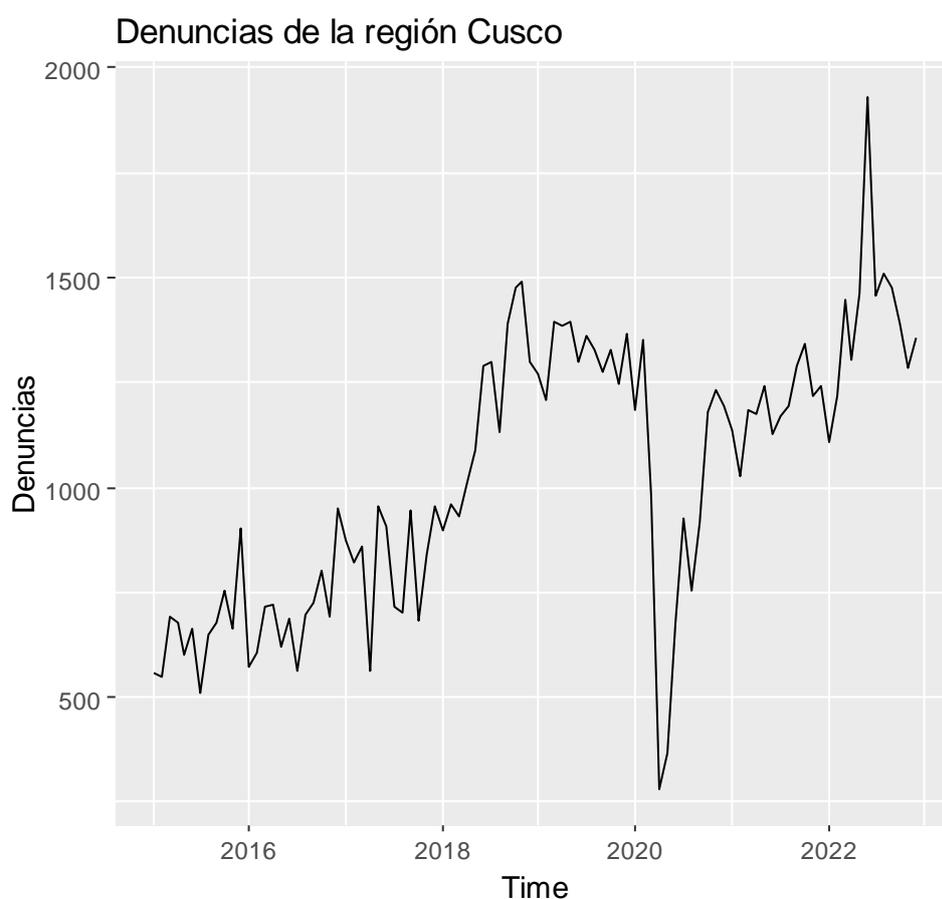
Con la finalidad de aplicar redes neuronales en la predicción de series temporales de las denuncias policiales de la región del Cusco, 2015-2022, se han analizado la cantidad de denuncias policiales mensuales que se han registrado de las distintas dependencias que están consideradas dentro de la región del Cusco.

4.1. ANÁLISIS DE LA COMPONENTE DE TENDENCIA EN LAS DENUNCIAS POLICIALES EN LA REGIÓN DEL CUSCO, 2015-2022

```
denuncias<-ts(datadenuncias$denuncias,start = 2015,frequency = 12)
autoplot(denuncias)
```

Figura 15

Gráfico de la serie de tiempo de las denuncias



La figura 15, representa la serie de tiempo de denuncias policiales en la región del Cusco revela patrones y tendencias significativas dignas de un análisis en

profundidad. Durante el período estudiado se observa una tendencia aparente de crecimiento sostenido en el número de denuncias presentadas a lo largo de los años previos a la aparición de la pandemia por COVID-19.

Sin embargo, el impacto de la pandemia en esta serie de tiempo es notable. A partir de la declaración de la emergencia sanitaria y las medidas de confinamiento, se registra una disminución abrupta y significativa en la cantidad de denuncias policiales presentadas en la región del Cusco. Este fenómeno puede ser atribuido a múltiples factores, incluyendo las restricciones de movilidad, el cierre de establecimientos comerciales y la reorientación de los recursos policiales hacia el cumplimiento de las medidas sanitarias, lo que podría haber reducido la capacidad de la población para presentar denuncias.

La variación en la serie de tiempo es particularmente interesante, ya que sugiere una marcada diferencia en la dinámica del sistema de denuncias antes y durante la pandemia. Este fenómeno puede ser analizado a través de técnicas estadísticas más avanzadas, como modelos de series temporales, para comprender mejor la relación entre la evolución de la pandemia y la cantidad de denuncias presentadas.

En resumen, la serie de tiempo de denuncias policiales en el Cusco muestra una tendencia de crecimiento antes de la pandemia, seguida de una reducción brusca durante el período de confinamiento por COVID-19. Esta dinámica es digna de un estudio más detallado para comprender las causas subyacentes y sus implicaciones en términos de seguridad pública en la región.

4.1.1. ANÁLISIS DE LA TENDENCIA: MANN KENDALL

Ho: No está presente el componente de tendencia

Ha: Esta presente el componente de tendencia

```
library(Kendall)
MannKendall(denuncias)
## tau = 0.544, 2-sided pvalue =< 2.22e-16
```

Como el pvalue es menor a 0.05, se rechaza la hipótesis nula, en tanto se puede asegurar mediante la prueba de Mann Kendall que la serie de tiempo de las denuncias en la región del Cusco presenta tendencia creciente.

4.1.2. ANÁLISIS DE LA TENDENCIA: COX STUART

Ho: No está presente el componente de tendencia

Ha: Esta presente el componente de tendencia

```
## Pruebas de tendencia: Cox Stuart
library(randtests)
cox.stuart.test(denuncias, "two.sided")

## Cox Stuart test
##
## data:  denuncias
## statistic = 43, n = 48, p-value = 1.368e-08
## alternative hypothesis: non randomness
```

Como el pvalue es menor a 0.05, se rechaza la hipótesis nula, en tanto se puede asegurar mediante la prueba de Cox Stuart que la serie de tiempo de las denuncias en la región del Cusco presenta tendencia.

4.1.3. ANÁLISIS DE LA TENDENCIA: CORRELACIÓN DE SPEARMAN

Ho: No está presente el componente de tendencia

Ha: Esta presente el componente de tendencia

```
cor.test(denuncias, time(denuncias), method="spearman")

## Spearman's rank correlation rho
## data:  denuncias and time(denuncias)
## S = 41466, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.7187588
```

Como el pvalue es menor a 0.05, se rechaza la hipótesis nula, en tanto se puede asegurar mediante la correlación de Spearman que la serie de tiempo y su el tiempo de la ocurrencia del hecho de las denuncias en la región del Cusco presenta tendencia.

4.1.4. ANÁLISIS DE LA TENDENCIA: WALD WOLFOWITZ

Ho: No está presente el componente de tendencia

Ha: Esta presente el componente de tendencia

```
## Wald-Wolfowitz test for independence and stationarity
## data:  denuncias
## z = 8.2479, n = 96, p-value < 2.2e-16
## alternative hypothesis: The series is significantly different from
## independence and stationarity
```

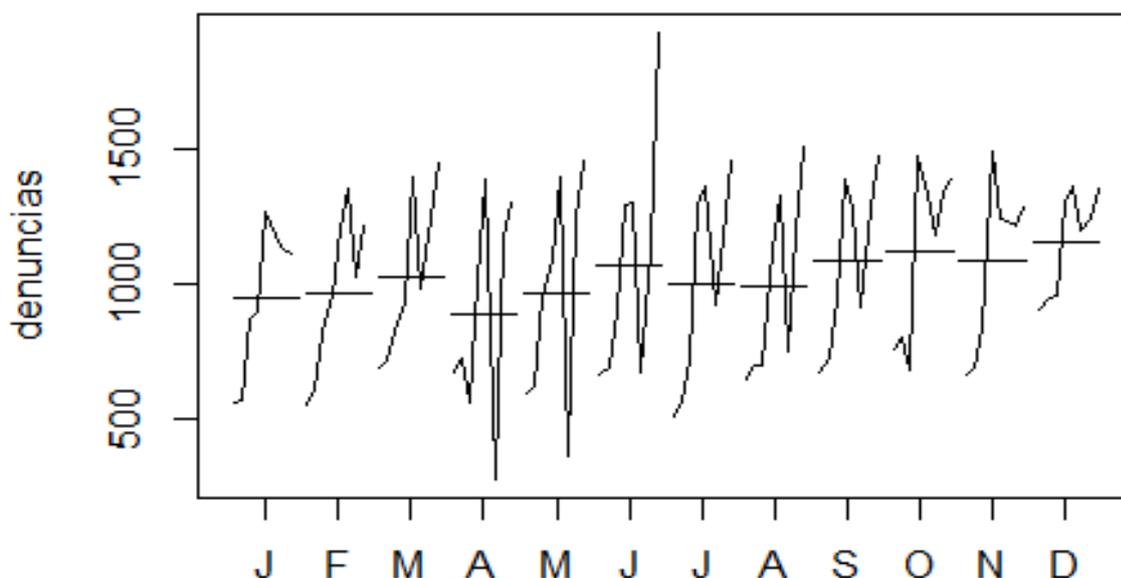
Como el pvalue es menor a 0.05, se rechaza la hipótesis nula, en tanto se puede asegurar mediante la prueba de Wald Wolfowitz que la serie de tiempo de las denuncias en la región del Cusco presenta tendencia.

4.2. ANÁLISIS DE LA COMPONENTE ESTACIONALIDAD EN LAS DENUNCIAS POLICIALES EN LA REGIÓN DEL CUSCO, 2015-2022

4.2.1. ANÁLISIS DE LA ESTACIONALIDAD: MONTHPLOT

Figura 16.

Monthplot



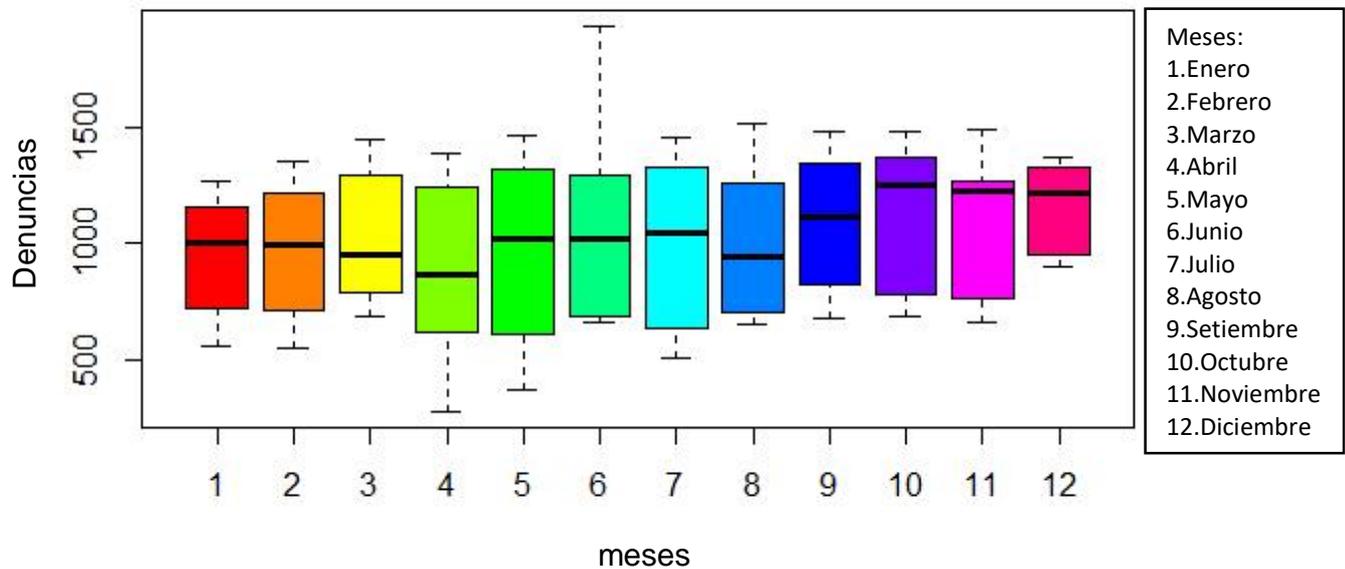
Nota. La figura en las abscisas presenta los meses J (January-enero), F (february-febrero) y así hasta llegar a D (december-diciembre)

La figura 16 muestra la ocurrencia de denuncias por cada mes, se observa que en promedio son similares, además en cuanto al tema de dispersión también se puede indicar que presenta similar dispersión con lo cual no se puede asegurar que haya estacionalidad, debido a que no existe diferencias significativas aparentes tanto en promedio como en dispersión.

4.2.2. ANÁLISIS de LA ESTACIONALIDAD: BOXPLOT

De la figura 17, se evidencia la ocurrencia de denuncias por cada mes, donde se observa que en promedio son similares, además en cuanto al tema de dispersión; también se puede indicar que presenta similar dispersión con lo cual no se puede asegurar que haya estacionalidad, debido a que no existe diferencias significativas aparentes tanto en promedio como en dispersión.

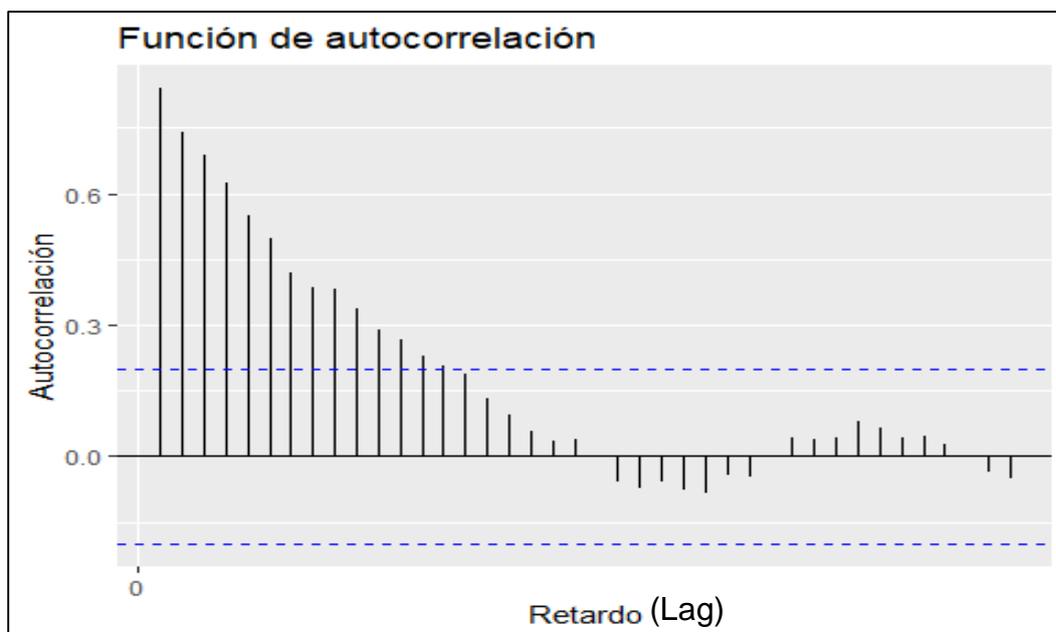
Figura 17
Gráfico de Cajas por meses



4.2.3. ANÁLISIS de la ESTACIONALIDAD: USO DEL CORRELOGRAMA

```
autoplot(TSA::acf(denuncias, lag = 40, plot = FALSE),
  xlab = "Retardo",
  ylab = "Autocorrelación",
  main = "Función de autocorrelación") + geom_hline(yintercept = 0)
```

Figura 18.
Función de autocorrelación



Por la función de autocorrelación y el gráfico 18, se observa que no existe un patrón cíclico, por lo que podemos afirmar que no existe estacionalidad en la serie de datos.

4.2.4. ANÁLISIS DE LA ESTACIONALIDAD: PRUEBA DE KRUSKAL WALLIS

Ho: No existe diferencias en la cantidad de denuncias en los diferentes meses. (no existe estacionalidad)

Ha: Existe diferencias en la cantidad de denuncias en los diferentes meses. (existe estacionalidad)

```
kruskal.test(denuncias ~ cycle(denuncias))
```

```
## Kruskal-Wallis rank sum test
##
## data:  denuncias by cycle(denuncias)
## Kruskal-Wallis chi-squared = 5.7274, df = 11, p-value = 0.8909
```

Como pvalue es 0.8909 no existe prueba suficiente para afirmar que existe estacionalidad en la serie de datos de denuncias policiales en la región del Cusco.

4.3. METODOLOGÍA BOX JENQUINS – ARIMA(p,d,q)

4.3.1. FASE DE IDENTIFICACIÓN

El modelo ARIMA (Autoregressive Integrated Moving Average) es una herramienta muy útil y poderosa en el análisis de series de tiempo que se utiliza para predecir valores futuros basados en patrones históricos. La identificación de la estacionariedad es un paso fundamental en la aplicación del modelo ARIMA, ya que ARIMA requiere que la serie de tiempo sea estacionaria o que se pueda convertir en estacionaria mediante diferenciación, esta se refiere a la propiedad estadística de la serie no cambian con el tiempo. Es decir, la media, la varianza y la autocorrelación de la serie deben permanecer constantes en el tiempo. Esto es esencial para que ARIMA funcione correctamente, ya que se basa en patrones repetitivos y predecibles.

La identificación de la estacionariedad implica visualizar la serie de tiempo y buscar señales de tendencia y estacionalidad. Para lograr la estacionariedad, a menudo es necesario diferenciar la serie de tiempo, es decir, calcular las diferencias entre valores sucesivos. Esto puede hacerse una o más veces hasta que la serie resultante sea estacionaria. Una vez que se ha logrado la estacionariedad, se pueden seleccionar los parámetros adecuados para el modelo ARIMA, como el orden de autoregresión (p) y el orden de media móvil (q). Estos parámetros determinarán cómo el modelo captura la estructura de la serie de tiempo.

```

# Determinando el valor de d
library(tseries)

kpss.test(denuncias) # si p_valor menor a 0.10 no es estacionario

## KPSS Test for Level Stationarity
## data:  denuncias
## KPSS Level = 1.4895, Truncation lag parameter = 3, p-value = 0.01

kpss.test(diff(denuncias))# conseguimos que el p_valor sea mayor que 0.10
por tanto es estacionaria

## KPSS Test for Level Stationarity
## data:  diff(denuncias)
## KPSS Level = 0.035979, Truncation lag parameter = 3, p-value = 0.1

# pruebas para determinar el número de diferencias tanto en la parte esta
cional como en la parte no estacional.
nsdiffs(denuncias)

## [1] 0

ndiffs(denuncias)

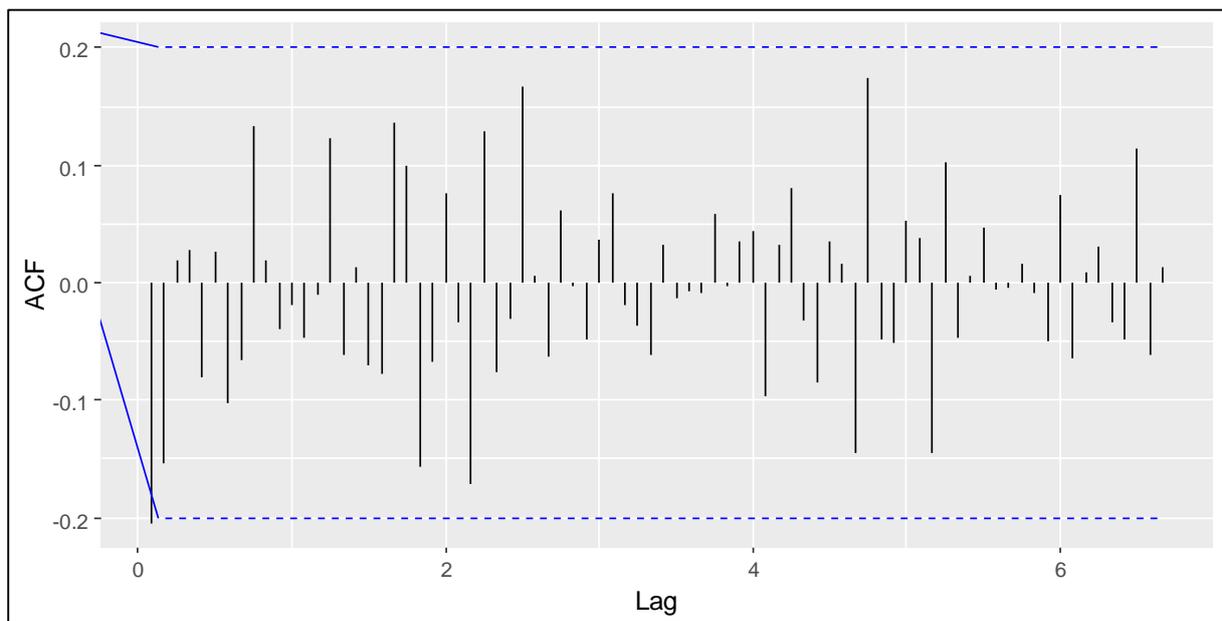
## [1] 1

```

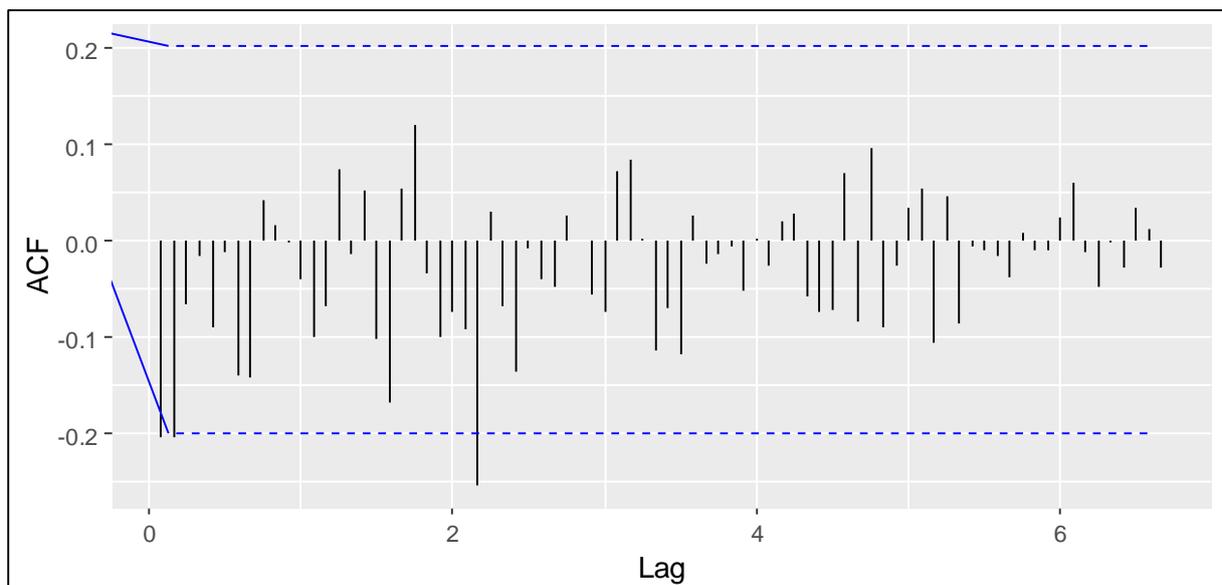
Después de haber diferenciado la serie de tiempo se consiguió probar estacionariedad con la prueba KPSS (Se utilizó para verificar si una serie de tiempo tiene una tendencia estacionaria en lugar de una tendencia que persiste en el tiempo), con ello se determinó que la serie de tiempo requiere de una diferencia para llegar a ser estacionaria.

Determinando el valor de p y q

```
acf(diff(denuncias), type = "correlation")
```

Figura 19. Autocorrelación

```
acf(diff(denuncias), type = "partial", main = "parcial")
```

Figura 20. Autocorrelación parcial

En base a las figuras 19 y 20 de autocorrelación y autocorrelación parcial se pudo determinar que como con una diferenciación se logró mantener constante la función de autocorrelación el valor de p debería ser 1, y observando la autocorrelación parcial se identificó irregularidad de no aleatoriedad por tanto se podría considerar q como 1.

4.3.2. FASE DE ESTIMACIÓN DE LOS MODELOS

Se considerarán 3 modelos a priori que podrían modelar el comportamiento de nuestros datos de las denuncias en la región del Cusco.

Modelo 1: ARIMA (1,1,0)

Modelo 2: ARIMA (1,1,1)

Modelo 3: ARIMA (2,1,0)

```
library(forecast)
mod1 = Arima(denuncias, order = c(1,1,0))
summary(mod1)

## Series: denuncias
## ARIMA(1,1,0)
##
## Coefficients:
##      ar1
##      -0.2003
## s.e.   0.1000
##
## sigma^2 = 28632: log likelihood = -621.78
## AIC=1247.55  AICc=1247.68  BIC=1252.66
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 9.803921 167.4392 120.4405 -2.140203 14.75571 0.4740195
##              ACF1
## Training set -0.04557323

mod2 = Arima(denuncias, order = c(1,1,1))
summary(mod2)

## Series: denuncias
## ARIMA(1,1,1)
##
## Coefficients:
##      ar1      ma1
##      0.5031 -0.7697
## s.e.   0.2459  0.1894
##
## sigma^2 = 27495: log likelihood = -619.43
## AIC=1244.85  AICc=1245.12  BIC=1252.52
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 18.34865 163.2043 116.7466 -1.759403 14.3721 0.4594814 -0.
01637905
```

```

mod3 = Arima(denuncias, order = c(2,1,0))
summary(mod3)

## Series: denuncias
## ARIMA(2,1,0)
##
## Coefficients:
##          ar1      ar2
##      -0.2402  -0.198
## s.e.   0.1002   0.100
##
## sigma^2 = 27772: log likelihood = -619.86
## AIC=1245.72  AICc=1245.98  BIC=1253.38
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
ACF1
## Training set 11.79618 164.0249 118.9941 -2.123811 14.53974 0.468327 -0.
01940494

auto.arima(denuncias, trace = TRUE, allowmean = TRUE)

##
## ARIMA(2,1,2)(1,0,1)[12] with drift      : Inf
## ARIMA(0,1,0) with drift                : 1251.383
## ARIMA(1,1,0)(1,0,0)[12] with drift     : 1251.495
## ARIMA(0,1,1)(0,0,1)[12] with drift     : 1249.138
## ARIMA(0,1,0)                          : 1249.522
## ARIMA(0,1,1) with drift                : 1247.174
## ARIMA(0,1,1)(1,0,0)[12] with drift     : 1249.127
## ARIMA(0,1,1)(1,0,1)[12] with drift     : 1251.347
## ARIMA(1,1,1) with drift                : Inf
## ARIMA(0,1,2) with drift                : 1246.608
## ARIMA(0,1,2)(1,0,0)[12] with drift     : 1248.699
## ARIMA(0,1,2)(0,0,1)[12] with drift     : 1248.692
## ARIMA(0,1,2)(1,0,1)[12] with drift     : 1250.964
## ARIMA(1,1,2) with drift                : Inf
## ARIMA(0,1,3) with drift                : 1248.712
## ARIMA(1,1,3) with drift                : Inf
## ARIMA(0,1,2)                          : 1245.234
## ARIMA(0,1,2)(1,0,0)[12]                : 1247.336
## ARIMA(0,1,2)(0,0,1)[12]                : 1247.335
## ARIMA(0,1,2)(1,0,1)[12]                : 1249.564
## ARIMA(0,1,1)                          : 1245.528
## ARIMA(1,1,2)                          : 1247.291
## ARIMA(0,1,3)                          : 1247.382
## ARIMA(1,1,1)                          : 1245.117
## ARIMA(1,1,1)(1,0,0)[12]                : 1247.235
## ARIMA(1,1,1)(0,0,1)[12]                : 1247.237
## ARIMA(1,1,1)(1,0,1)[12]                : Inf
## ARIMA(1,1,0)                          : 1247.682
## ARIMA(2,1,1)                          : 1247.297
## ARIMA(2,1,0)                          : 1245.981
## ARIMA(2,1,2)                          : 1249.443

```

```
##
## Best model: ARIMA(1,1,1)

## Series: denuncias
## ARIMA(1,1,1)
##
## Coefficients:
##          ar1          ma1
##      0.5031  -0.7697
## s.e.  0.2459   0.1894
##
## sigma^2 = 27495: log likelihood = -619.43
## AIC=1244.85  AICc=1245.12  BIC=1252.52

mod4 = Arima(denuncias, order = c(1,1,1))
```

En base a los anteriores modelos, se encontró el modelo que presenta menor error a la hora de ajustarse el cual es el *Modelo 2: ARIMA (1,1,1)*, veamos la siguiente comparación, con los modelos mas cercanos al modelo, los cuales tienen un mejor desempeño en predicción de nuevos casos, se comparara los indicadores de bondad de ajuste o también conocido como indicadores de validación.

Tabla 3. Comparación de modelos ARIMA

	ARIMA (1,1,0) <i>Modelo 1</i>	ARIMA (1,1,1) <i>Modelo 2</i>	ARIMA (2,1,0) <i>Modelo 3</i>
AIC	1247.55	1244.85	1245.72
ME	9.8039	18.3486	11.7962
RMSE	167.4392	163.2043	164.0249
MAE	120.4405	116.7466	118.9941
MPE	-2.1402	-1.7594	-2.1238
MAPE	14.7557	14.3721	14.5397

Observando los indicadores de precisión como el AIC, RMSE, MAE, MPE, MAPE en todos ellos el modelo 2, es el modelo que presenta menor AIC, RMSE, MAE, MPE, MAPE que son indicadores de error, por ello consideraremos el modelo 2 como el mejor modelo ARIMA para poder hacer las estimaciones y predicciones de la serie de datos.

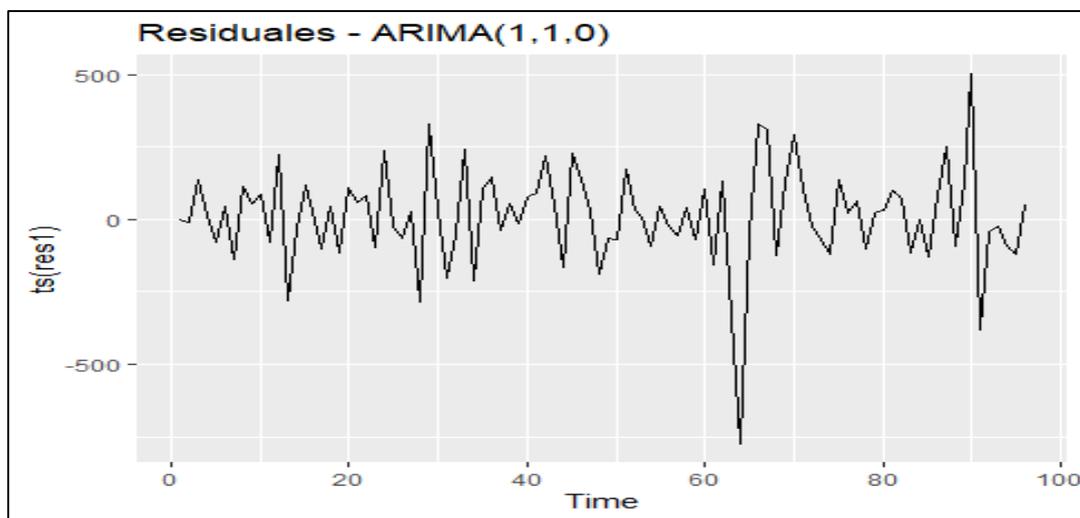
4.3.3. FASE DE DIAGNÓSTICO Y VALIDACIÓN

Para esta etapa consideraremos los residuos para ver algunos supuestos, en principio se observa a continuación el correlograma de los residuos obtenido de los modelos propuestos para poder competir según la metodología del ARIMA.

```
res1 = residuals(mod1)
res2 = residuals(mod4)

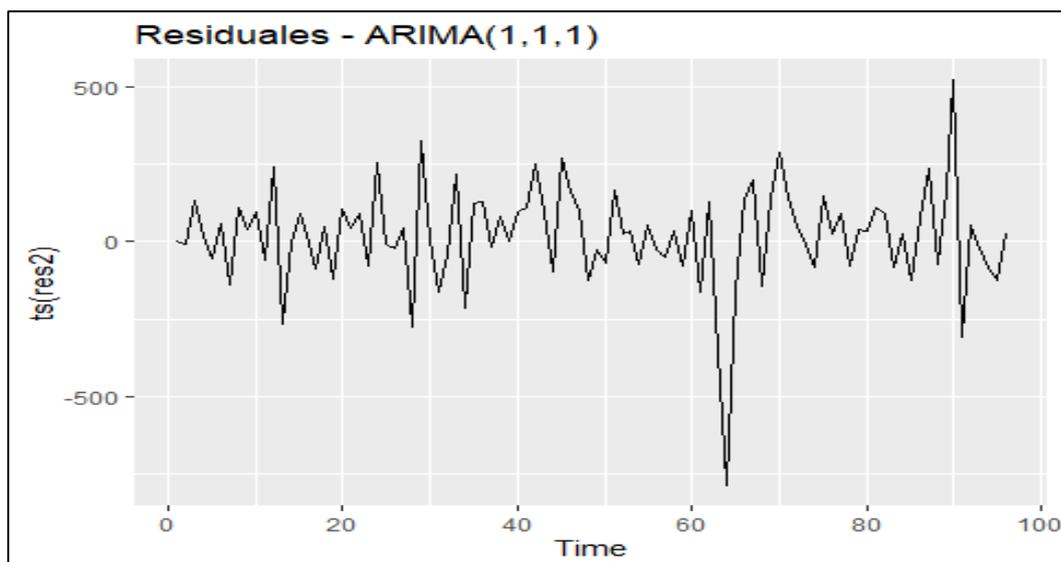
autoplot(ts(res1),
          main = "Residuales - ARIMA(1,1,0)")
```

Figura 21. Residuales ARIMA (1,1,0)



```
autoplot(ts(res2),
          main = "Residuales - ARIMA(1,1,1)")
```

Figura 22. Residuales ARIMA(1,1,1)



En las figuras 21 y 22 se observan una alta volatilidad, debido a que en la época de pandemia paso a bruscamente a bajar la cantidad de denuncias.

A continuación, se realizará una prueba estadística para verificar si la media está centrada en Cero.

$$H_0: \mu_r = 0$$

$$H_a: \mu_r \neq 0$$

```
t.test(res)

## One Sample t-test
## t = 0.57168, df = 95, p-value = 0.5689
## mean of x
## 9.803921

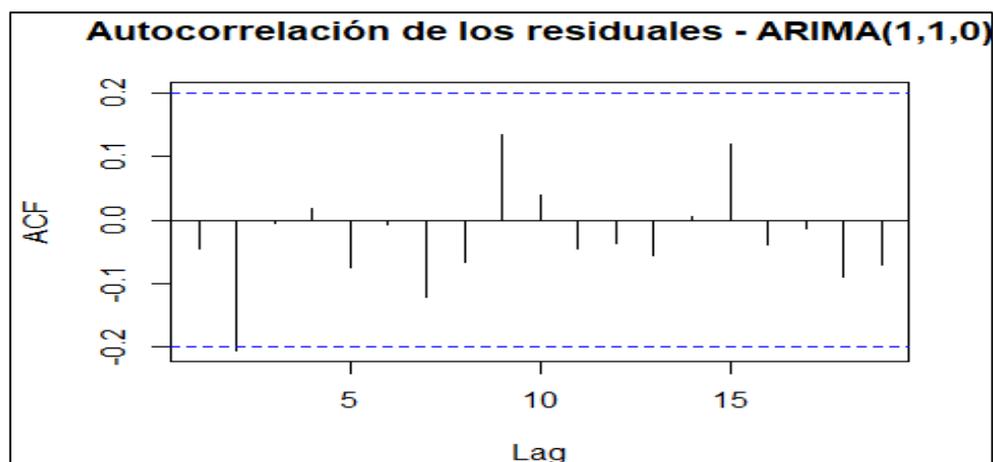
t.test(res2)

## One Sample t-test
## t = 1.1028, df = 95, p-value = 0.2729
## mean of x
## 18.34865
```

Se observa que los $pvalue > 0.05$, por tanto, no existe evidencia estadística para rechazar la hipótesis de que la media sea cero.

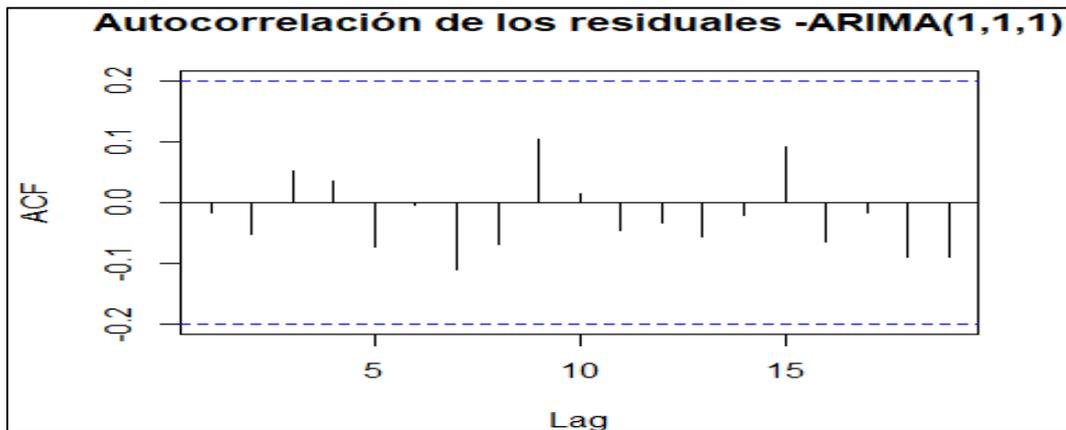
```
TSA::acf(ts(res1),
  main = "Autocorrelación de los residuales - ARIMA(1,1,0)")
```

Figura 23. Autocorrelación de los residuales- ARIMA (1,1,0)



```
TSA::acf(ts(res2),
  main = "Autocorrelación de los residuales -ARIMA(1,1,1)")
```

Figura 24. Autocorrelación de los residuales-ARIMA(1,1,1)



Se observa en los gráficos 23 Y 24 de autocorrelación, los residuos no tienen dependencia con sus valores pasados en los distintos rezagos, existe independencia en los residuos de ambos modelos. En el gráfico 23 un ligero rezago en el lag 2, pero lo consideraremos dentro de lo normal.

Validación del modelo de serie temporal ARIMA 1,1,0 y ARIMA 1,1,1

Se ha segmentado los datos, los datos correspondientes al 2022 son datos los cuales tomaremos como datos de prueba, donde se ha observado las predicciones que realizaron los modelos considerados y se calcularon las medidas de precisión correspondientes:

```
# Validación
denuncias.train = window(denuncias, start = 2015, end = 2021, frequency = 12)
denuncias.test  = window(denuncias, start = 2022)
mod1.train      = Arima(denuncias.train, order = c(1,1,0))
accuracy(forecast(mod1.train,18),denuncias.test)

##                ME      RMSE      MAE      MPE      MAPE      MASE
## Training set   9.188314 168.5508 121.2915 -2.762994 16.57896 0.5042446
## Test set      275.375713 367.8082 284.8194 17.249822 18.10137 1.1840784
##                ACF1 Theil's U
## Training set  -0.03279389      NA
## Test set      0.21712741  1.400982

mod2.train      = Arima(denuncias.train, order = c(1,1,1))
accuracy(forecast(mod2.train,18),denuncias.test)

##                ME      RMSE      MAE      MPE      MAPE      MASE
## Training set  16.44207 164.2041 117.4561 -2.506055 16.18783 0.4882996
## Test set     346.56877 423.7855 346.5688 22.407414 22.40741 1.4407889
##                ACF1 Theil's U
## Training set  -0.0121971      NA
## Test set     0.2173501  1.623758
```

Para el año 2022, el modelo que presenta mejores indicadores de precisión es el modelo1 ARIMA (1,1,0), es decir el nivel de error de las predicciones con respecto a la serie de datos original es menor al modelo 2, por tanto, como el modelo 1 se acerca más a la serie de datos original en el último tramo de la serie, será con este modelo que se hará las predicciones correspondientes para el 2023.

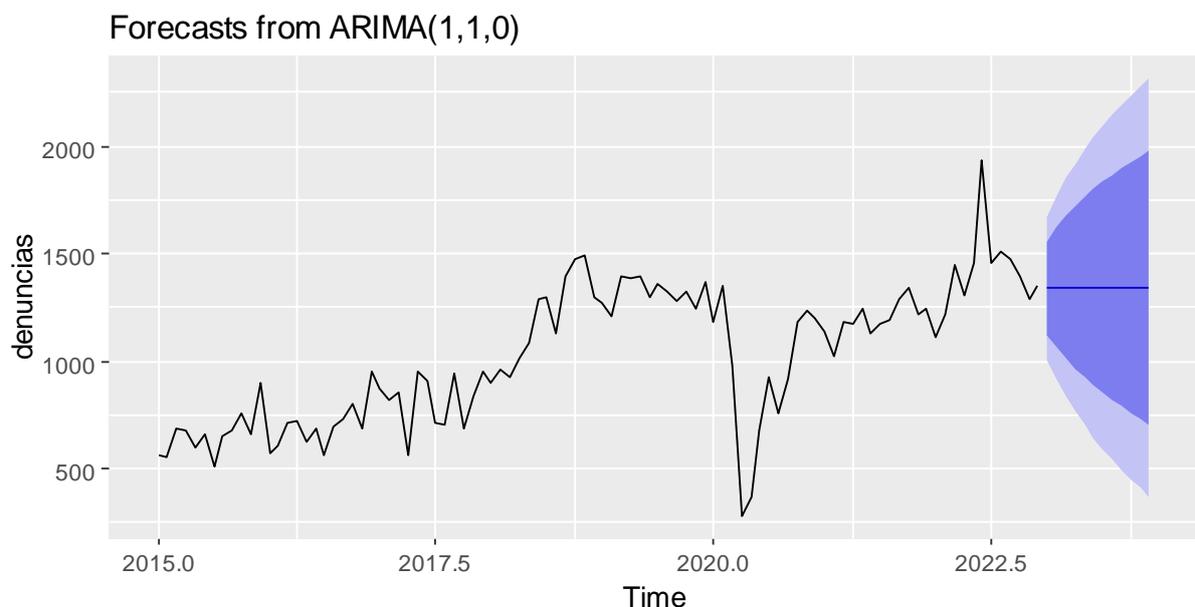
4.3.4. FASE DE PREDICCIÓN ARIMA (1,1,0)

```
forecast(mod1, h = 12)
```

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## Jan 2023	1340.178	1123.3250	1557.030	1008.5301	1671.825
## Feb 2023	1342.947	1065.2835	1620.610	918.2975	1767.596
## Mar 2023	1342.392	1010.3341	1674.450	834.5534	1850.230
## Apr 2023	1342.503	964.6194	1720.387	764.5799	1920.426
## May 2023	1342.481	923.6063	1761.355	701.8676	1983.094
## Jun 2023	1342.485	886.3161	1798.654	644.8347	2040.136
## Jul 2023	1342.484	851.8421	1833.127	592.1117	2092.857
## Aug 2023	1342.485	819.6381	1865.331	542.8599	2142.109
## Sep 2023	1342.484	789.3054	1895.664	496.4701	2188.499
## Oct 2023	1342.485	760.5517	1924.417	452.4950	2232.474
## Nov 2023	1342.485	733.1533	1951.816	410.5928	2274.376
## Dec 2023	1342.485	706.9349	1978.034	370.4953	2314.474

```
plot(forecast(mod1, h = 12))
```

Figura 25. Predicción con el modelo ARIMA(1,1,0)



4.4. APLICACIÓN DE REDES NEURONALES EN SERIES TEMPORALES DE LAS DENUNCIAS POLICIALES EN LA PROVINCIA DEL CUSCO, 2015-2022.

4.4.1. FASE DE AJUSTE DEL MODELO

```
# Ajuste del modelo
library(forecast)
library(ggplot2)
fit = nnetar(denuncias)
fit

## Series: denuncias
## Model: NNAR(1,1,2)[12]
## Call: nnetar(y = denuncias)
##
## Average of 20 networks, each of which is
## a 2-2-1 network with 9 weights
## options were - linear output units
##
## sigma^2 estimated as 20545
```

Para datos no estacionales, el modelo ajustado se denota como un modelo NNAR(p , k), donde k es el número de nodos ocultos. Esto es análogo a un modelo AR(p) pero con funciones no lineales. Para datos estacionales, el modelo ajustado se llama un modelo NNAR(p , P , k)[m], que es análogo a un modelo ARIMA(p , 0, 0)(P , 0, 0)[m] pero con funciones no lineales.

De acuerdo a los datos obtenidos del ajuste del modelo se obtiene que la red neuronal que se utilizó fue el NNAR(1,1,2)[12] el cual fue un promedio de 20 redes, cada una de las cuales es una red 2-2-1 con 9 opciones de pesos, siendo las unidades de salida lineales, y se podría indicar que es un ARIMA(1, 0, 0)(1, 0, 0)[12] con dos capas ocultas ($k=2$) considerando funciones no lineales en las capas ocultas.

4.4.2. PRECISIÓN Y VALIDACIÓN DEL MODELO

```
accuracy(forecast(fit))

##           ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
## Training set 0.05019413 147.2228 105.375 -3.766999 12.69475 0.4147259 0.1584902
```

Tabla 4. Comparación de Precisión y validación de los modelos ARIMA Y Red Neuronal

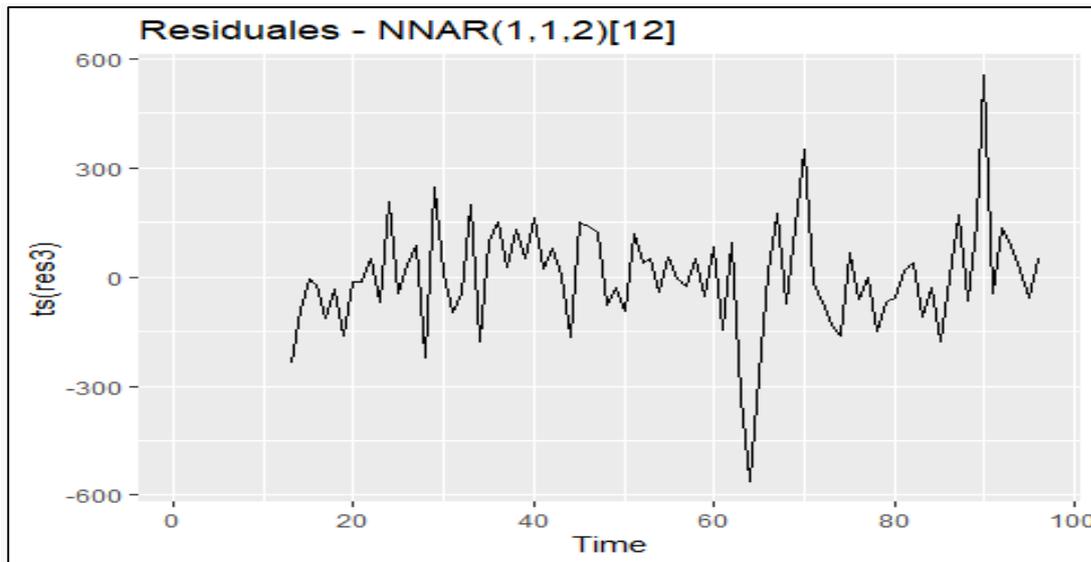
	<i>ARIMA (1,1,0)</i> <i>Modelo 1</i>	<i>ARIMA (1,1,1)</i> <i>Modelo 2</i>	<i>NNAR (1,1,2)[12]</i> <i>Modelo red Neuronal</i>
<i>AIC</i>	<i>1247.55</i>	<i>1244.85</i>	<i>---</i>
<i>ME</i>	<i>9.8039</i>	<i>18.3486</i>	<i>0.0502</i>
<i>RMSE</i>	<i>167.4392</i>	<i>163.2043</i>	<i>147.2228</i>
<i>MAE</i>	<i>120.4405</i>	<i>116.7466</i>	<i>105.375</i>
<i>MPE</i>	<i>-2.1402</i>	<i>-1.7594</i>	<i>-3.766999</i>
<i>MAPE</i>	<i>14.7557</i>	<i>14.3721</i>	<i>12.69475</i>

Se puede observar que en la tabla 4 el error medio (ME) del modelo que considera la red neuronal presenta menor error (0.0502), así mismo en la raíz del cuadrado medio del error (RMSE) es menor a los considerados con los modelos ARIMAs, lo mismo sucede con el error medio absoluto (MAE =105.375), MAPE (Error Porcentual Absoluto Medio) también es menor con las redes neuronales. Por consiguiente el modelo NNAR (1,1,2)[12] presenta mejores indicadores de precisión.

```
res3 = residuals(fit)
autoplot(ts(res3),
         main = "Residuales - NNAR(1,1,2)[12]")
```

Graficando los residuales que se han generado utilizando la red neuronal en series de tiempo, se visualiza que en promedio está centrado en cero, se podría indicar que la serie de datos presenta estacionariedad.

Figura 26. Residuales - NNAR(1,1,2)[12]



```
t.test(res3)
```

```
## One Sample t-test
## t = 0.0031061, df = 83, p-value = 0.9975
## mean of x
## 0.05019413
```

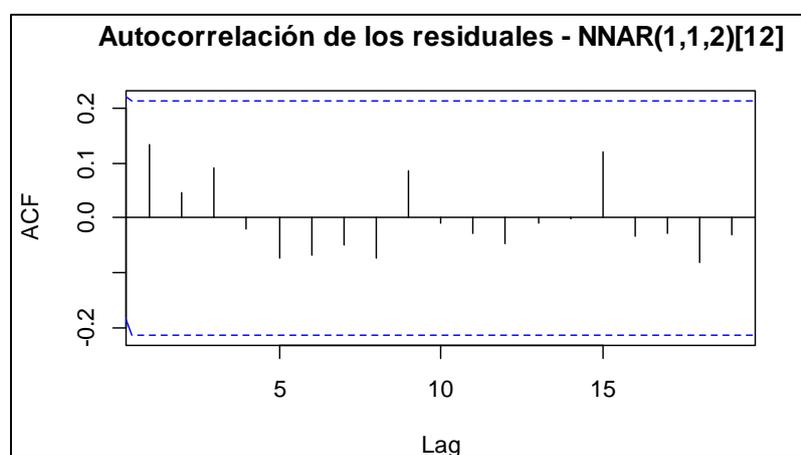
$H_0: \mu_r = 0$

$H_a: \mu_r \neq 0$

De la prueba t, se confirma con un 5% de significancia que en promedio los residuales obtenidos considerando las redes neuronales están centradas en cero (p-value=0.9975).

```
TSA::acf(ts(res3[-c(1:12)]),
  main = "Autocorrelación de los residuales - NNAR(1,1,2)[12]")
```

Figura 27. Autocorrelación de los residuales- NNAR(1,1,2)[12]



En la figura 27 se puede evidenciar que los residuales, no presentan autocorrelación significativa, por tanto, se podría indicar que los residuos tienen un comportamiento sin dependencia en el tiempo por los desfases,

5.4.3. FASE DE PREDICCIÓN NNAR (1,1,2) [12]

Se utilizo los tres últimos meses del año 2023, considerando la red neuronal para series de tiempo NNAR (1,1,2) [12], indicando que para los 3 últimos meses del año 2023 se espera unas estimaciones de 1299.55, 1302.46 y 1302.403 denuncias respectivamente.

Predicciones

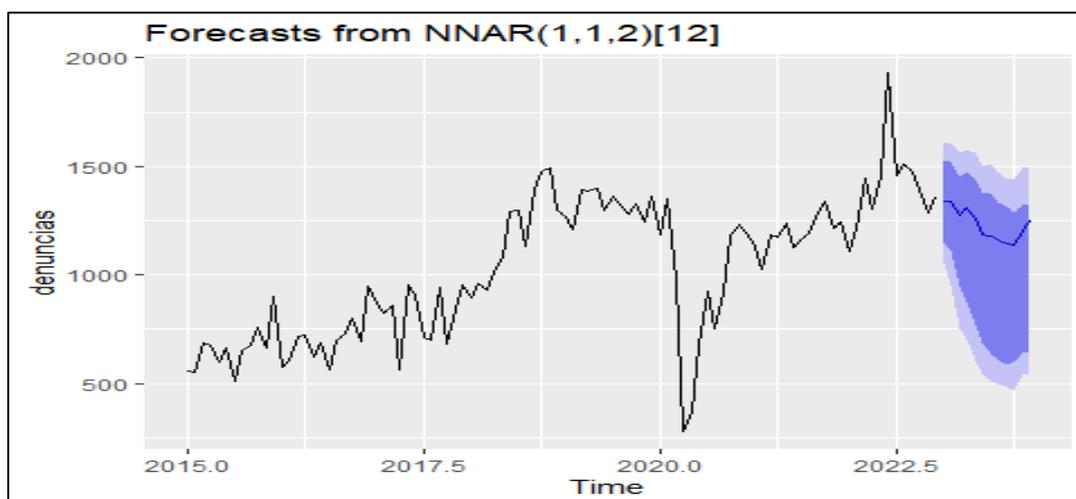
```
nnetforecast <- forecast(fit, h = 12, PI = T)
nnetforecast
```

##	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
## Jan 2023	1333.563	1139.8990	1516.175	1044.5843	1612.623
## Feb 2023	1323.735	1090.8062	1521.896	952.7473	1623.567
## Mar 2023	1309.205	997.7513	1508.630	754.4498	1626.509
## Apr 2023	1305.640	922.6945	1504.421	687.5144	1601.720
## May 2023	1301.821	803.0822	1479.058	627.4589	1576.724
## Jun 2023	1280.408	741.1713	1471.634	580.0676	1593.679
## Jul 2023	1291.079	695.0372	1455.402	548.8736	1555.553
## Aug 2023	1294.354	676.1199	1453.069	536.0532	1580.951
## Sep 2023	1296.742	655.5072	1440.387	533.3933	1577.369
## Oct 2023	1299.550	673.5578	1454.036	546.2824	1578.080
## Nov 2023	1302.459	671.9855	1447.322	510.6538	1548.382
## Dec 2023	1302.403	664.7734	1433.444	519.5274	1562.411

```
library(ggplot2)
```

```
autoplot(nnetforecast)
```

Figura 28. Proyecciones con el modelo NNAR(1,1,2)[12]



4.5. DISCUSIÓN DE RESULTADOS

Fernández, (2021) quiso demostrar la utilidad de las redes neuronales en el contexto de análisis de series de tiempo, y en la presente investigación se confirma que las redes neuronales recurrente (Long Short-Term Memory) a la hora de ajustarse a series temporales en este caso particular a denuncias policiales tienen un mejor ajuste a comparación de los modelos de series de tiempo ARIMA, este mismo concepto de que esta red neuronal recurrente tiene muy buen ajuste lo comparten con (López & Sales, 2020) también.

Agustín, (2017) logro predecir movimientos a 15 días de una compañía del índice S&P 500, EBAY, aplicando redes neuronales y LSTM (long short time memory) a partir de únicamente los históricos de los precios de cierre de 500 compañías, además destacar que los movimientos se han clasificado en tres clases siendo 1 comprar, 0 mantenerse y -1 vender. Esto nos confirma que las redes neuronales tienen un alto desempeño tanto para predecir como para clasificar, al igual que en el presente estudio donde demostramos que las redes neuronales recurrentes tienen un mejor performance a la hora predecir cantidad de denuncias en la Región del Cusco.

(Fierro, 2020) demostró que los patrones predictivos del modelo estadístico parecen suavizarse o desaparecer a medida que se extiende el horizonte. Para el caso de los modelos de aprendizaje automático (redes neuronales) esto no se observa, los patrones que aprenden los modelos se mantienen. Esto puede explicar intuitivamente el motivo por el cual los modelos de redes neuronales responden mejor para horizontes mayores, y este resultado no es ajeno a lo encontrado en la presente tesis debido a que las predicciones con los modelos ARIMA son constantes en el tiempo, y con el modelo de redes neuronales el patrón tiene variaciones propias de la serie de datos de denuncias de los delitos cometidos en la región del Cusco.

(Saavedra, 2017) en su investigación llega a concluir que ha identificado un patrón en la ejecución presupuestaria durante el período comprendido entre 2006 y 2016 y que a nivel nacional, se observa un aumento progresivo y constante en la ejecución presupuestaria en ese lapso temporal, coincidiendo con un aumento correlativo en las denuncias por delitos, las infracciones y los casos de violencia familiar, tanto de carácter físico como psicológico, son delitos que han experimentado un incremento

durante dicho período, la tendencia que ellos lo comprobaron de manera visual nosotros la comprobamos con técnicas estadísticas para verificar tendencia creciente en la presente investigación, confirmando lo expresado por (Saavedra, 2017).

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

- Con relación a la componente de la tendencia se puede asegurar mediante las pruebas de Mann Kendall, Cox Stuart, Wald Wolfowitz y correlación de Spearman la serie de tiempo de las denuncias policiales en la región del Cusco presenta tendencia creciente, lo cual nos indica que las políticas del gobierno para disminuir la violencia familiar, inseguridad ciudadana, y demás no están siendo efectivas, debido a la tendencia creciente de las denuncias policiales a como era antes de la pandemia por el Covid-19.
- Con respecto a la componente de estacionalidad se pudo comprobar que de la observación del grafico de monthplot y la gráfica de cajas, así como de la prueba de kruskall wallis ($pvalue = 0.8909$) no existen pruebas suficientes para afirmar que existe estacionalidad en la serie de datos de denuncias policiales en la región del Cusco, en tanto no habría indicios para afirmar que en algunos meses del año haya más denuncias policiales o más delitos
- En base a los indicadores de precisión como el AIC, RMSE, MAE, MPE, MAPE los modelo ARIMA (1,1,1) y el modelo (1,1,0), son los modelos que presenta menor indicio de error a comparación de los demás modelos, por ello consideraremos estos dos modelos para poder hacer las estimaciones, evaluaciones y validaciones, así como las predicciones de la serie de datos.
- La red neuronal que se utilizo fue el NNAR(1,1,2)[12] el cual fue un promedio de 20 redes, cada una de las cuales es una red 2-2-1 con 9 opciones de pesos, siendo las unidades de salida lineales, y se podría indicar que es un ARIMA(1, 0, 0)(1, 0, 0)[12] con dos capas ocultas ($k=2$) considerando funciones no lineales en las capas ocultas. Su error medio ($ME=0.0502$) el cual es el menor error a comparación de los modelos ARIMAs; así mismo en la raíz del cuadrado medio del error (RMSE), el error medio absoluto ($MAE =105.375$), MAPE (Error Porcentual Absoluto Medio) también es menor con las redes

neuronales. Para el año 2023, considerando la red neuronal para series de tiempo NNAR (1,1,2) [12], para los 3 últimos meses del año 2023 se espera unas estimaciones de 1299.55, 1302.46 y 1302.403 denuncias respectivamente.

RECOMENDACIONES

Evaluar Políticas de Seguridad Ciudadana, dado que se ha identificado una tendencia creciente en las denuncias policiales en la región del Cusco, es importante profundizar en la evaluación de las políticas de seguridad ciudadana implementadas por el gobierno local y regional. Verificar si estas políticas están alineadas con las necesidades reales de la comunidad y si están teniendo un impacto positivo en la reducción de la violencia familiar y la inseguridad ciudadana.

En futuras investigaciones evaluar las causas de la tendencia creciente, buscar posibles causas detrás de la tendencia creciente en las denuncias policiales. Se podría propiciar implicar la realización de encuestas o entrevistas a la población para comprender mejor por qué están aumentando las denuncias. Además, se puede investigar si la percepción de seguridad de la población ha cambiado y si eso ha influido en el aumento de las denuncias.

Aunque no se encontró evidencia de estacionalidad en los datos, aún es importante realizar un análisis temporal más detallado. Se propone dividir los datos en subconjuntos, por ejemplo, antes y después de la pandemia por COVID-19, y analizar si hay patrones estacionales en cada uno de estos subconjuntos.

Analizar el posible impacto socioeconómico de un aumento en las denuncias policiales en la región del Cusco. ¿Cómo afecta esto a la percepción de seguridad de la población? ¿Tiene repercusiones en el turismo o en la inversión en la región? Esto puede proporcionar información valiosa para la toma de decisiones gubernamentales. Así mismo se propone la exploración de otras variables que podrían estar relacionadas con las denuncias policiales con las redes neuronales recurrentes (LSTM) y comparar estas acciones con las otras regiones del Perú.

REFERENCIAS BIBLIOGRÁFICAS

- Acosta-Cervantes, M., Villareal-Marroquin, M., & Cabrera-Rios, M. (2013). Estudio de validación de un método para seleccionar técnicas de pronóstico de series de tiempo mediante redes neuronales artificiales. *Ingeniería, Investigación y Tecnología*, 53-63. doi:[https://doi.org/10.1016/S1405-7743\(13\)72225-X](https://doi.org/10.1016/S1405-7743(13)72225-X).
- Agustín, P. (2017). *Redes Neuronales Recurrentes: Una aplicación para los mercados bursátiles*. España: Universitat de Barcelona y Universitat Politecnica de Catalunya.
- Anderson, D. (1985). *Times Series Analysis and Forecasting*. Londres: Editions Lonchers Polytechnics y Butterworths.
- Basoqain, X. (sf). *Redes neuronales artificiales y sus aplicaciones*. Ehu: Escuela superior de Ingenieria de Bilbao.
- BCRP. (2015). *CARACTERIZACIÓN DEL DEPARTAMENTO DE CUSCO*. Cusco: BCRP .
- Bertona, L. (2005). *Entrenamiento de redes neuronales basado en algoritmos evolutivos*. Buenos aires: Universidad de buenos aires.
- Bowerman, B., & Koehler, A. (2007). *Pronósticos, Series de Tiempo Y Regresión: Un Enfoque Aplicado*. International Thomson Editores, S.A.
- Cabello, E. (2004). *Técnicas de reconocimiento facial mediante redes neuronales*. Madrid: Universidad Politecnica de Madrid. doi:10.20868/UPM.thesis.215
- Calvo, A. (2009). *Predicción en series de tiempo con Modelos Aditivos*. España: Universidade da Coruña.
- Canavos, G. (1988). *Probabilidad y estadística, aplicaciones y métodos*. México: McGraw-Hill.
- Cano, N. (15 de abril de 2023). *El Peruano diario oficial*. Obtenido de Inseguridad ciudadana como preocupación : https://www.vanderbilt.edu/lapop/news/081115_elperuano.pdf
- Ccopa, O. (2019). *Aplicación móvil para la gestión de denuncias de los efectivos policiales en la Policía Nacional del Perú*. Andahuaylas: Universidad Nacional José María Arguedas.
- Crespo, w. (2016). *Análisis de la variabilidad espacial y temporal de la precipitación en la cuenca del río Pampas, de 1975 al 2014*. Ayacucho: Universidad Nacional de San Cristóbal de Huamanga.
- De la Torre, C. (2007). *Combinación de clasificadores en redes neurales*. Lima: Universidad Nacional Mayor de San Marcos.
- Defensoría del Pueblo. (2013). *¡No Callar! Guía práctica para la denuncia ciudadana contra actos de corrupción y otras faltas contra la ética pública*. Lima, Perú: Defensoría del Pueblo.

- Fernandez, R. (2021). *Series Temporales Avanzadas: Aplicación de Redes Neuronales para el Pronóstico de Series de Tiempo*. España: Universidad de Granada.
- Fierro, A. (2020). *Predicción de Series Temporales con Redes Neuronales*. Argentina: Universidad Nacional del la Plata.
- Gulli, A., & Pal, S. (2017). *Deep Learning with Keras: Implementing deep learning models and neural networks with the power of Python*. Birmingham-Mumbai: Packt.
- INEI. (2018). *Victimización en el Perú 2010-2017*. Lima: Instituto Nacional de Estadística e Informática del Perú.
- INEI. (2021). *Informe técnico: estadísticas de la criminalidad, seguridad ciudadana y violencia*. Lima: INEI. Obtenido de https://www.inei.gov.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1892/libro.pdf
- INEI. (2023a). *Victimización en el Perú 2015-2022*. Lima: INEI. Obtenido de https://www.inei.gov.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1913/libro.pdf
- INEI. (2023b). *Informe técnico n°5 agosto 2023: Estadísticas de seguridad ciudadana*. Lima: INEI. Obtenido de <https://m.inei.gov.pe/media/MenuRecursivo/boletines/estadisticas-de-seguridad-ciudadana-enero-2023-junio-2023.pdf>
- Ley N°30506. (2017). *Ley que delega en el Poder Ejecutivo la Facultad de legislar en materia de reactivación económica y formalización, seguridad ciudadana, lucha contra la corrupción, agua y saneamiento y reorganización de Petroperú S.A*. Lima, Perú: Presidencia de la república.
- López, J., & Sales, O. (2020). *Implementación hardware de una red neuronal Long Short-Term Memory*. Madrid: Universidad Complutense.
- Martín del Brio, B., & Sanz, A. (2006). *Redes Neuronales y Sistemas Borrosos*. España: Universidad de Zaragoza.
- Martínez, E. (2004). *Métodos no paramétricos I (Notas de clase)*. Buenos Aires, Argentina: instituto de calculo - departamento de matematica - FCEYN – UBA. Obtenido de https://www.dm.uba.ar/materias/optativas/metodos_no_parametricos/2004/2/
- Match, D. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*. Rosario: Universidad Tecnológica Nacional – Facultad Regional Rosario.
- Maya, M. (2022). *Pronóstico de series de tiempo empleando redes neuronales y meta-transferencia de aprendizaje*. México: Centro de investigación y de estudios avanzados del instituto politécnico nacional.
- Ministerio público . (20 de Enero de 2024). *Cómo denunciar*. Obtenido de *Cómo se formula una denuncia*: https://www.mpfm.gob.pe/fiscalias_anticorrupcion/como_denunciar/

- Muñoz, A. (1996). *Aplicación de técnicas de redes neuronales artificiales al diagnóstico de procesos industriales*. Madrid: Universidad Pontificia Comillas Madrid.
- Namigtle, A. (2018). *Diagnóstico de Fallas en un Riel de Inyectores*. Morelos, México: Tecnológico Nacional de México.
- Núñez, Y. (2015). *Inseguridad ciudadana y delito en el Perú*. Pimentel: Universidad Señor de Sipan.
- Peña, D. (2000). *Modelos lineales y Series Temporales*. Madrid: Edición Alianza Editorial.
- Quispe, A., Calla, K., Yangali, J., Rodríguez, J., & Pumacayo, I. (2019). *Estadística no paramétrica aplicada a la investigación científica con software SPSS, MINITAB y EXCEL*. Colombia: Eidec editorial.
- Ramos, F. (2017). *Modelos Univariantes para describir y predecir la serie de nacimientos y defunciones del distrito de Acora, periodo 1994-2015*. Puno: Universidad Nacional del Altiplano.
- Rejas, A., Diaz, E., & Troya, H. (2022). La evaluación del presupuesto por resultados y su impacto en la seguridad ciudadana del Perú. *Ciencia Latina Revista Científica Multidisciplinar*, 2995-3011. doi:https://doi.org/10.37811/cl_rcm.v6i1.1701
- Reyna, C., & Toche, E. (1999). *La Inseguridad en el Perú*. Lima: CEPAL. División de Desarrollo Social.
- Rubio, J. (20 de Enero de 2024). *IA Delta*. Obtenido de Inteligencia Artificial, Regresión lineal y descenso de gradiente: <https://iadelta.com/inteligencia-artificial/regresion/lineal-y-gradiente/>
- Saavedra, M. (2017). Inversión pública en seguridad ciudadana y los indicadores de delitos, faltas y violencia familiar en el Perú, 2006 - 2016. *Revista Gobierno y gestión pública 2017*, 97-116. Obtenido de <https://www.aulavirtualusmp.pe/ojs/index.php/RevistaGobiernoYG/article/view/2324>
- Santamaría, F. (2019). *Estudio y predicción de activos financieros mediante redes neuronales*. España: Universidad Politecnica Madrid.
- Tablada, C. (sf). *Redes Neuronales Artificiales*. Argentina: Universidad Nacional de Córdoba. Obtenido de <https://www.famaf.unc.edu.ar/~revm/digital24-3/redes.pdf>
- UNIR. (20 de Enero de 2024). *La universidad en Internet*. Obtenido de ¿Qué es el algoritmo backpropagation para el entrenamiento de redes neuronales?: <https://www.unir.net/ingenieria/revista/backpropagation/>
- Uriel Jimenez, E. (2012). *Análisis de Series temporales, modelos Arima*. Madrid: Editorial Paraniunfo S.A.
- Vorobioff, J., Cerrotta, S., Morel, N., & Amadio, A. (2022). *Inteligencia Artificial y Redes Neuronales: Fundamentos, Ejercicios y Aplicaciones con Python y Matlab*. Buenos Aires: Universidad Tecnológica Nacional Buenos Aires.

ANEXOS

Anexo 1. MATRIZ DE CONSISTENCIA

Título: REDES NEURONALES EN LA PREDICCIÓN DE SERIES TEMPORALES DE LAS DENUNCIAS POLICIALES DE LA REGIÓN DEL CUSCO, 2015-2022

Tabla 5. Matriz de consistencia

PROBLEMA GENERAL	OBJETIVO GENERAL	HIPOTESIS GENERAL	VARIABLE:
<p>¿Cuáles son los modelos de redes neuronales que mejor predicen series temporales de las denuncias policiales de la región del Cusco, 2015-2022?</p> <p>PROBLEMAS ESPECÍFICOS</p> <p>a) ¿Cuál es la evolución temporal de las denuncias policiales en la región del Cusco durante el período comprendido entre 2015 y 2022, en su componente de la tendencia?</p> <p>b) ¿Cuál es el patrón estacional observado en las denuncias policiales en la región del Cusco durante el periodo de 2015 a 2022?</p>	<p>Determinar los modelos de redes neuronales en series temporales que mejor predice a la serie de datos de las denuncias policiales en la región del Cusco, 2015-2022.</p> <p>OBJETIVOS ESPECIFICOS</p> <p>a) Analizar la evolución temporal de las denuncias policiales en la región del Cusco durante el período 2015-2022, con respecto de la componente tendencia.</p> <p>b) Examinar el patrón estacional presente en las denuncias policiales en la región del Cusco durante</p>	<p>Las redes neuronales recurrente multicapa, al emplearse en la predicción de series temporales de denuncias policiales en la Región del Cusco para el período de 2015 a 2022, demuestren un desempeño superior en comparación con el modelo ARIMA</p> <p>HIPÓTESIS ESPECÍFICOS</p> <p>a) Existe variabilidad significativa en la naturaleza y evolución temporal de las denuncias policiales en la región del Cusco entre 2015 y 2022.</p> <p>b) Existe un patrón estacional en las denuncias policiales en la región del Cusco durante el periodo 2015-2022.</p>	<p>DENUNCIAS POLICIALES</p>

<p>c) ¿Cuál es el modelo ARIMA que mejor ajusta a la serie de datos de las denuncias policiales en la provincia del Cusco, 2015-2022, para la predicción de los siguientes meses?</p> <p>d) ¿Cuál es el modelo de redes neuronales en series temporales que mejor ajusta a la serie de datos de las denuncias policiales en la provincia del Cusco, 2015-2022?</p>	<p>el periodo comprendido entre 2015 y 2022.</p> <p>c) Determinar el modelo ARIMA que mejor se ajusta a la serie temporal de datos de las denuncias policiales en la provincia del Cusco durante el periodo de 2015 a 2022.</p> <p>d) Determinar el modelo de redes neuronales en series temporales que mejor se ajusta a la serie de datos de las denuncias policiales en la provincia del Cusco durante el periodo de 2015 a 2022</p>	<p>c) Existe un modelo ARIMA específico que ofrece un ajuste óptimo a la serie temporal de denuncias policiales en la provincia del Cusco.</p> <p>d) Las redes neuronales LSTM en series temporales ofrecerá una predicción que se ajusta a la serie de datos de denuncias policiales en la provincia del Cusco</p>	
--	---	---	--

Anexo 2. DATOS PROPORCIONADOS PNP

MES		AÑO	MACRO REGION	DIRECCION POLICIAL / FRENTE P	NOMBRE DE DEPARTAMENTO	NOMBRE DE LA PROVINCIA	NOMBRE DEL DISTRITO	NOMBRE DE LA COMISARIA	TOTAL DENUNCIAS DELITOS	TOTAL DELITOS CONTRA LA VIDA, EL CUERPO Y LA	I. DELITOS CONTRA LA VIDA, EL CUERPO Y LA SALUD														TOTAL LESIONES LEVES											
										TOTAL HOMICIDIOS	I.1. HOMICIDIOS										I.2. ABORTO															
											HOMICIDIO SIMPLE	HICIDIO	HICIDIO CALIFICADO (ASESINATO)	HICIDIO	ARIATO	CONVENIENCIA Y EL CRECIMIENTO DEL DELITO DE SECUESTRO	HICIDIO POR EMOCION VIOLENTA	ANTICIDIO	HICIDIO CULPOSO	HICIDIO PASOSO	IGACION O AYUDA AL SUICIDIO	HICIDIO POR PAF	FATIVA	TOTAL ABORTO	ABORTO	ORTO CONSENTIDO	ORTO SIN CONSENTIMIENTO	ORTO PRETERINTENCIONAL	ORTO TERAPEUTICO	ORTO SENTIMENTAL	ORTO EMBRIGUICO	ORTO CON MUERTE SUBITA	FATIVA			
237	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	CUSCO	COMISARIA TAHUANTINSUYO	64	5	0													0											0	
238	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	CUSCO	COMISARIA SECTORIAL CUSCO	34	3	0													0											3	1
239	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	CUSCO	COMISARIA SIPASPUCYO	20	3	0													0											3	2
240	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	POROY	COMISARIA POROY	17	2	1	1												0											1	1
241	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SAN JERONIMO	COMISARIA SAN JERONIMO	55	2	0													0											2	
242	MARZO	2018	DIRECCION DE TRANSITO TRANSPORTE Y SEGURIDAD VIAL	TRANS	CUSCO	CUSCO	SAN JERONIMO	COMISARIA PROTECCION CARRETERAS CUSCO	0	0	0													0											0	
243	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SAN SEBASTIAN	COMISARIA SAN SEBASTIAN	83	4	0													0											4	3
244	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SANTIAGO	COMISARIA LA FAMILIA	19	18	0													0											18	
245	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SANTIAGO	COMISARIA INDEPENDENCIA	22	0	0													0											0	
246	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SANTIAGO	COMISARIA VIVA EL PERU	44	20	0													1		1									19	2
247	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SANTIAGO	COMISARIA ZARZUELA	30	1	0													0											1	
248	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SANTIAGO	COMISARIA SANTIAGO	38	2	0													0											2	
249	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	SAYLLA	COMISARIA SAYLLA	7	2	0													0											2	2
250	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	WANCHAQ	COMISARIA DEL AEROPUERTO ALEJANDRO VEL	12	1	1										1			0											0	
251	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	CUSCO	WANCHAQ	COMISARIA SECTORIAL WANCHAQ	102	11	0													1		1									10	
252	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	ACOMAYO	ACOMAYO	COMISARIA ACOMAYO	0	0	0													0												
253	MARZO	2018	CUSCO- APURIMAC (VII)	POLICIA	CUSCO	ACOMAYO	ACOPIA	COMISARIA ACOPIA	1	1	0													0												

Anexo 3. FICHA DE RECOLECCIÓN DE DATOS

La presente ficha servirá para poder generar la data consolidada de todas las denuncias que han sido registradas en el sistema de denuncias de la Policía Nacional del Perú.

Año	Mes	Nro. Denuncias
2015	Enero	
2015	Febrero	
2015	Marzo	
.	.	.
.	.	.
.	.	.
2022	Noviembre	
2022	Diciembre	

Anexo 4. CÓDIGO DE R

```

library(forecast)
library(ggplot2)
library(readxl)
library(readxl)
datadenuncias <- read_excel("dataFinal.xlsx")

##### PRUEBAS DE TENDENCIA#####
## pruebas de tendencia: Mann Kendall
library(Kendall)
MannKendall(denuncias)

## Pruebas de tendencia: Correlaci?n de Spearman
cor.test(denuncias,time(denuncias),method="spearman")

##### PRUEBAS DE ESTACIONALIDAD #####

## Pruebas de estacionalidad: Uso de boxplot
boxplot(denuncias~cycle(denuncias),
        xlab="meses",
        ylab = "Denuncias",
        col=rainbow(12))
autoplot(TSA::acf(denuncias, lag = 40, plot = FALSE),
        xlab = "Retardo",
        ylab = "Autocorrelaci?n",
        main = "Funci?n de autocorrelaci?n") +
geom_hline(yintercept = 0)

## Pruebas de estacionalidad: Prueba de Kruskal Wallis
kruskal.test(denuncias ~ cycle(denuncias))

```

```
##### diferenciación #####
# Determinando el valor de d
library(tseries)
kpss.test(denuncias) #
kpss.test(diff(denuncias))#

# Determinando el valor de p y q
library(TSA)
autoplot(TSA::acf(diff(denuncias),type="correlation",lag =
80,plot=FALSE))
autoplot(TSA::acf(diff(denuncias),type="partial",lag =
80,plot=FALSE))

#####
#### estimacion #####
library(forecast)
mod1 = Arima(denuncias, order = c(1,1,0))
summary(mod1)

mod2 = Arima(denuncias, order = c(1,1,1))
summary(mod2)

mod3 = Arima(denuncias, order = c(2,1,0))
summary(mod3)

auto.arima(denuncias, trace = TRUE, allowmean = TRUE)

mod4 = Arima(denuncias, order = c(1,1,1))
#### diagnostico####
# Análisis de residuales
```

```
res1 = residuals(mod1)
res2 = residuals(mod4)

autoplot(ts(res1),
          main = "Residuales - ARIMA(1,1,0)")
autoplot(ts(res2),
          main = "Residuales - ARIMA(1,1,1)")

t.test(res1)
t.test(res2)

TSA::acf(ts(res1),
          main = "Autocorrelación de los residuales - ARIMA(1,1,0)")
TSA::acf(ts(res2),
          main = "Autocorrelación de los residuales -ARIMA(1,1,1)")

# medidas de precision
accuracy(forecast(mod1))
accuracy(forecast(mod2))

# Validación
denuncias.train = window(denuncias, start = 2015, end = 2021,
frequency = 12)
denuncias.test  = window(denuncias, start = 2022)
mod1.train  = Arima(denuncias.train, order = c(1,1,0))
accuracy(forecast(mod1.train,18),denuncias.test)
```

```
mod2.train = Arima(denuncias.train, order = c(1,1,1))
accuracy(forecast(mod2.train,18),denuncias.test)
```

```
##### paso 4 Predicción
```

```
forecast(mod1, h = 12)
autoplot(forecast(mod1, h = 12))
```

```
#####
```

```
##      redes neuronales en series de tiempo
```

```
# Ajuste del modelo
```

```
library(forecast)
```

```
library(ggplot2)
```

```
fit = nnetar(denuncias)
```

```
fit
```

```
# precision y validacion del modelo
```

```
accuracy(forecast(fit))
```

```
res3 = residuals(fit)
```

```
autoplot(ts(res3),
```

```
          main = "Residuales - NNAR(1,1,2)[12]")
```

```
t.test(res3)
```

```
TSA::acf(ts(res3[-c(1:12)]),
```

```
          main = "Autocorrelación de los residuales -
NNAR(1,1,2)[12]")
```

```
# Predicciones
```

```
nnetforecast <- forecast(fit, h = 24, PI = T)
```

```
nnetforecast
```

```
library(ggplot2)
```

```
autoplot(nnetforecast)
```

Anexo 5. DATOS

Denuncias policiales 2015-2022								
mes	2015	2016	2017	2018	2019	2020	2021	2022
1	559	573	874	895	1270	1184	1134	1109
2	550	605	823	960	1207	1350	1027	1216
3	690	717	858	930	1393	980	1184	1445
4	676	722	565	1011	1386	279	1175	1305
5	600	621	954	1087	1396	368	1240	1460
6	662	688	905	1291	1300	678	1127	1931
7	509	561	714	1297	1362	927	1169	1456
8	651	697	702	1131	1327	754	1193	1510
9	676	727	945	1390	1277	917	1288	1475
10	756	802	684	1475	1326	1178	1342	1390
11	663	691	840	1491	1246	1231	1216	1285
12	903	949	955	1298	1365	1196	1242	1354

Anexo 6. SOLICITUDES DE INFORMACIÓN

SOLICITA: INFORMACION ESTADISTICA PARA
CONSOLIDAR MI TRABAJO DE INVESTIGACION

SEÑORA:

MAYOR SHIANDY GARY MERCADO ATAYUPANQUI
JEFE DE UNITIC DE LA VII MACRO REGION POLICIAL CUSCO
Ciudad.-

JESSICA HUANCACHOQUE COHAILA, identificada con DNI 41614781, con domicilio en Av. Los Incas 1006-B, Distrito de Wanchaq , provincia y departamento del Cusco, numero de celular 984583030 y con correo electrónico jheska@gmail.com, anta usted con el debido respeto me presento y expongo lo siguientes:

Que solicito a su representada información estadística para consolidar mi trabajo de investigación intitulada "REDES NEURONALES EN LA PREDICCIÓN DE SERIES TEMPORALES DE LAS DENUNCIAS POLICIALES DE LA REGIÓN DEL CUSCO, 2015-2022", para la Universidad Nacional San Antonio Abad del Cusco información requerida a:

Cantidad de denuncias por delito de la provincia Cusco de los años de 2015 al 2022 por años, meses, semanas.

Solicitud que me permita requerir, con el fin de concretar mi trabajo de investigación y en amparo del texto único ordenado de la ley N° 27806 Ley de transparencia y Acceso a la información pública.

POR LO EXPUESTO: Ruego a Ud. Acceder a lo solicitado, tomando en consideración que el fin es educativo y de formación profesional.

Cusco, 23 de enero de 2023

Adjunto:
Copia de Dni



JESSICA HUANCACHOQUE COHAILA
DNI N° 41614781





ESTADISTICA VII MACREPOL <estadistica7cusco@gmail.com>

REMITE INFORMACION DE DELITOS 2015 AL 2022

1 mensaje

ESTADISTICA VII MACREPOL <estadistica7cusco@gmail.com>

27 de enero de 2023, 18:26

Para: "jheska@gmail.com" <jheska@gmail.com>

BUENAS TRADES SE REMITE LA INFORMACION SOLICITADA DE DELITOS SEGUN LO SOLICITADO.
ATTE.

MAY. S PNP SHIANDY GARY MERCADO ATAYUPANQUI
JEFE UNITIC VII MACREPOL CUSCO
CEL 931269805

VII MACRO REGIÓN POLICIAL CUSCO
OFICINA DE ESTADÍSTICA

Plaza Túpac Amaru S/N - Wanchaq - Cusco

Confidencialidad: Este correo es confidencial y va dirigido exclusivamente a su destinatario. Si lo ha recibido por error por favor comuníquelo al remitente y elimínelo de inmediato sin copiarlo, imprimirlo, utilizarlo o manipularlo de modo alguno.



DELITOS 2018 AL 2022 D13 JESICA HUANCACHOQUE.xlsx
12030K