



UNIVERSIDAD ESAN  
FACULTAD DE INGENIERÍA  
INGENIERÍA DE TECNOLOGÍAS DE INFORMACIÓN Y SISTEMAS

**Desarrollo de un Algoritmo de aprendizaje profundo para detectar rutas seguras de las  
calles de los Ángeles**

Trabajo de investigación para el curso de Trabajo de Tesis I

Nombre Guillermo Díaz Beltrán

Asesor: Marks Calderón

Lima, 28 de abril de 2024

## Resumen

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ac odio tempor orci dapibus ultrices in iaculis nunc sed. Vivamus arcu felis bibendum ut tristique et egestas quis ipsum. Odio morbi quis commodo odio aenean sed adipiscing diam donec. Donec ultrices tincidunt arcu non sodales neque sodales ut. Fusce ut placerat orci nulla pellentesque dignissima enim sit amet. Facilisi etiam dignissima diam quis enim lobortis. Sit amet justo donec enim diam vulputate ut pharetra. Gravida in fermentum et sollicitudin ac orci phasellus egestas. Ultricies tristique nulla aliquet enim tortor at auctor. Nullam vehicula ipsum a arcu cursus vitae congue mauris. Convallis posuere morbi leo urna molestie at elementum eu facilisis. Elit at imperdiet dui accumsan sit amet nulla. Amet consectetur adipiscing elit pellentesque habitant morbi tristique senectus et. Mauris in aliquam sem fringilla ut morbi. Ultricies integer quis auctor elit sed vulputate mi sit. Nulla pellentesque dignissima enim sit amet venenatis urna cursus eget. Ac feugiat sed lectus vestibulum mattis ullamcorper. Eu augue ut lectus arcu bibendum. Rhoncus dolor purus non enim praesent elementum.

Nulla facilisi cras fermentum odio eu feugiat pretium. Massa massa ultricies mi quis hendrerit. Id leo in vitae turpis massa sed elementum. Quis vel eros donec ac odio tempor orci. Netus et malesuada fames ac turpis egestas integer eget aliquet. Velit ut tortor pretium viverra suspendisse potenti. Ut enim blandit volutpat maecenas. Nibh tellus molestie nunc non blandit. Mus mauris vitae ultricies leo integer malesuada nunc vel. Vel elit scelerisque mauris pellentesque pulvinar pellentesque habitant. Neque viverra justo nec ultrices dui sapien eget. Vitae aliquet nec ullamcorper sit. Dui id ornare arcu odio ut sem nulla pharetra diam. Et magnis dis parturient montes. Varius morbi enim nunc faucibus.

**Palabras claves:** uno, dos, tres, cuatro

## Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ac odio tempor orci dapibus ultrices in iaculis nunc sed. Vivamus arcu felis bibendum ut tristique et egestas quis ipsum. Odio morbi quis commodo odio aenean sed adipiscing diam donec. Donec ultrices tincidunt arcu non sodales neque sodales ut. Fusce ut placerat orci nulla pellentesque dignissima enim sit amet. Faciliti etiam dignissima diam quis enim lobortis. Sit amet justo donec enim diam vulputate ut pharetra. Gravida in fermentum et sollicitudin ac orci phasellus egestas. Ultricies tristique nulla aliquet enim tortor at auctor. Nullam vehicula ipsum a arcu cursus vitae congue mauris. Convallis posuere morbi leo urna molestie at elementum eu facilisis. Elit at imperdiet dui accumsan sit amet nulla. Amet consectetur adipiscing elit pellentesque habitant morbi tristique senectus et. Mauris in aliquam sem fringilla ut morbi. Ultricies integer quis auctor elit sed vulputate mi sit. Nulla pellentesque dignissima enim sit amet venenatis urna cursus eget. Ac feugiat sed lectus vestibulum mattis ullamcorper. Eu augue ut lectus arcu bibendum. Rhoncus dolor purus non enim praesent elementum.

Nulla facilisi cras fermentum odio eu feugiat pretium. Massa massa ultricies mi quis hendrerit. Id leo in vitae turpis massa sed elementum. Quis vel eros donec ac odio tempor orci. Netus et malesuada fames ac turpis egestas integer eget aliquet. Velit ut tortor pretium viverra suspendisse potenti. Ut enim blandit volutpat maecenas. Nibh tellus molestie nunc non blandit. Mus mauris vitae ultricies leo integer malesuada nunc vel. Vel elit scelerisque mauris pellentesque pulvinar pellentesque habitant. Neque viverra justo nec ultrices dui sapien eget. Vitae aliquet nec ullamcorper sit. Dui id ornare arcu odio ut sem nulla pharetra diam. Et magnis dis parturient montes. Varius morbi enim nunc faucibus.

**Keywords:** uno, dos, tres, cuatro

Para mi X, Y,X

## Agradecimientos

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ac odio tempor orci dapibus ultrices in iaculis nunc sed. Vivamus arcu felis bibendum ut tristique et egestas quis ipsum. Odio morbi quis commodo odio aenean sed adipiscing diam donec. Donec ultrices tincidunt arcu non sodales neque sodales ut. Fusce ut placerat orci nulla pellentesque dignissima enim sit amet. Faciliti etiam dignissima diam quis enim lobortis. Sit amet justo donec enim diam vulputate ut pharetra. Gravida in fermentum et sollicitudin ac orci phasellus egestas. Ultricies tristique nulla aliquet enim tortor at auctor. Nullam vehicula ipsum a arcu cursus vitae congue mauris. Convallis posuere morbi leo urna molestie at elementum eu facilisis. Elit at imperdiet dui accumsan sit amet nulla. Amet consectetur adipiscing elit pellentesque habitant morbi tristique senectus et. Mauris in aliquam sem fringilla ut morbi. Ultricies integer quis auctor elit sed vulputate mi sit. Nulla pellentesque dignissima enim sit amet venenatis urna cursus eget. Ac feugiat sed lectus vestibulum mattis ullamcorper. Eu augue ut lectus arcu bibendum. Rhoncus dolor purus non enim praesent elementum.

Nulla facilisi cras fermentum odio eu feugiat pretium. Massa massa ultricies mi quis hendrerit. Id leo in vitae turpis massa sed elementum. Quis vel eros donec ac odio tempor orci. Netus et malesuada fames ac turpis egestas integer eget aliquet. Velit ut tortor pretium viverra suspendisse potenti. Ut enim blandit volutpat maecenas. Nibh tellus molestie nunc non blandit. Mus mauris vitae ultricies leo integer malesuada nunc vel. Vel elit scelerisque mauris pellentesque pulvinar pellentesque habitant. Neque viverra justo nec ultrices dui sapien eget. Vitae aliquet nec ullamcorper sit. Dui id ornare arcu odio ut sem nulla pharetra diam. Et magnis dis parturient montes. Varius morbi enim nunc faucibus.

# Índice general

<b>Índice de Figuras</b>	<b>8</b>
<b>Índice de Tablas</b>	<b>9</b>
<b>1. PLANTEAMIENTO DEL PROBLEMA</b>	<b>10</b>
1.1. Descripción de la Realidad Problemática . . . . .	10
1.2. Formulación del Problema . . . . .	12
1.2.1. Problema General . . . . .	12
1.2.2. Problemas Específicos . . . . .	13
1.3. Objetivos de la Investigación . . . . .	13
1.3.1. Objetivo General . . . . .	13
1.3.2. Objetivos Específicos . . . . .	13
1.4. Justificación de la Investigación . . . . .	14
1.4.1. Teórica . . . . .	14
1.4.2. Práctica . . . . .	14
1.4.3. Metodológica . . . . .	15
1.5. Delimitación del Estudio . . . . .	15
1.5.1. Espacial . . . . .	15
1.5.2. Temporal . . . . .	15
1.5.3. Conceptual . . . . .	15

1.6.	Hipótesis . . . . .	16
1.6.1.	Hipótesis General . . . . .	16
1.6.2.	Hipótesis Específicas . . . . .	16
1.6.3.	Matriz de Consistencia . . . . .	16
<b>2.</b>	<b>MARCO TEÓRICO</b>	<b>17</b>
2.1.	Antecedentes de la investigación . . . . .	17
2.1.1.	Copper price estimation using bat algorithm (Dehghani & Bogdanovic, 2018) . . . . .	17
2.2.	Bases Teóricas . . . . .	18
2.2.1.	Machine Learning . . . . .	18
2.2.2.	Natural Language Processing (NLP) . . . . .	18
2.3.	Marco Conceptual . . . . .	19
<b>3.</b>	<b>METODOLOGÍA DE LA INVESTIGACIÓN</b>	<b>20</b>
3.1.	Diseño de la investigación . . . . .	20
3.1.1.	Diseño no experimental . . . . .	20
3.1.2.	Tipo explicativo . . . . .	20
3.1.3.	Enfoque cuantitativo . . . . .	21
3.2.	Población y muestra . . . . .	21
3.3.	Operacionalización de Variables . . . . .	21
3.4.	Instrumentos de medida . . . . .	22
3.5.	Técnicas de recolección de datos . . . . .	22
3.6.	Técnicas para el procesamiento y análisis de la información . . . . .	23
3.7.	Cronograma de actividades y presupuesto . . . . .	23
<b>4.</b>	<b>DESARROLLO DEL EXPERIMENTO</b>	<b>24</b>
4.1.	X . . . . .	24

4.2. Y . . . . .	24
4.3. Z . . . . .	25
<b>5. ANÁLISIS Y DISCUSIÓN DE RESULTADOS</b>	<b>26</b>
5.1. X . . . . .	26
5.2. Y . . . . .	26
5.3. Z . . . . .	27
<b>6. CONCLUSIONES Y RECOMENDACIONES</b>	<b>28</b>
6.1. Conclusiones . . . . .	28
6.2. Recomendaciones . . . . .	28
<b>Anexos</b>	<b>29</b>
<b>A. Anexo I: Matriz de Consistencia</b>	<b>30</b>
A.1. Árbol de Problemas . . . . .	30
A.2. Árbol de Objetivos . . . . .	32
<b>B. Anexo II: Resumen de Papers investigados</b>	<b>33</b>
<b>BIBLIOGRAFÍA</b>	<b>35</b>



# Índice de Figuras

1.1. Índice de criminalidad de los países de América 2023 . . . . .	11
1.2. Días con más robos de celulares registrados en el 2023 Osiptel . . . . .	12
3.1. Prueba de Figura . . . . .	21

# Índice de Tablas

3.1. An example table. . . . .	23
4.1. An example table. . . . .	24
5.1. An example table. . . . .	26
A.1. Matriz de consistencia. Fuente: Elaboración propia . . . . .	31
B.1. Cuadro Resumen de Papers investigados. Fuente: Elaboración propia . . . . .	34

# Capítulo 1

## PLANTEAMIENTO DEL PROBLEMA

### 1.1. Descripción de la Realidad Problemática

De acuerdo al análisis del crecimiento y expansión urbana del Centro Nacional de Planeamiento Estratégico (CEPLAN) del 2023, menciona que con respecto al censo 2017 del INEI, ha crecido las zonas urbanas a un 1.6% en comparación al censo de 2007, lo que quiere decir que la población urbana ha crecido en aproximadamente 3 millones de habitantes (CEPLAN, 2023). Este crecimiento de la urbanización también a traído consigo un aumento significativo a las acciones criminales que enfrenta Perú, y también en todo el mundo; como es el caso de California (Estados Unidos), ya que se revela que los crímenes violentos, robos y asaltos agravados ha aumentado en un 5.7% en el año 2023 (PPIC, 2023). Esto refleja una gran preocupación de la seguridad ciudadana en la que enfrenta la población y las autoridades que buscan tomar medidas cruciales para proteger a los residentes de los entornos urbanos.

El índice de criminalidad en Lima representa un 70.9%, que se sitúa en el puesto número 17, por debajo de ciudades de Latinoamérica como Caracas (Venezuela), San Pedro Sula (Honduras) y Río de Janeiro (Brasil) que también enfrentan varios desafíos para reducir las acciones delictivas y mejorar la seguridad ciudadana (Numbeo, 2023). A pesar de eso, el índice de representatividad de crímenes en Lima no resta importancia a la clara necesidad de encontrar alternativas efectivas que permita reducir dicho índice de criminalidad en otras áreas urbanas del Perú.

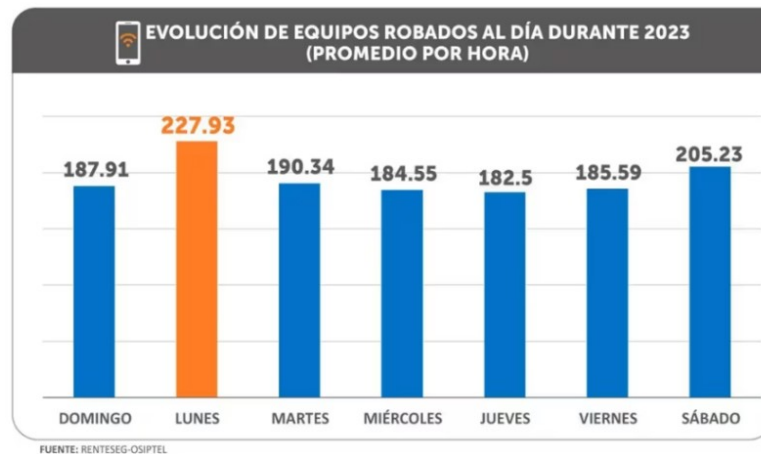
Según datos recopilados en la plataforma de Statista entre Julio y diciembre del 2023, por cada 100 habitantes en las zonas urbanas de Perú, el 11.9% de esta población ha sido víctimas por el robo de pertenencias personales, como dinero, carteras o celulares, siendo esta, la más representativa y la que refleja la parte vulnerable de los ciudadanos al no hacer frente

Puesto	Ciudad	Índice de Criminalidad	Índice de Seguridad
1	Caracas, Venezuela	83,59	16,41
2	San Pedro Sula, Honduras	80,55	19,45
3	Río de Janeiro, Brasil	77,63	22,37
4	Fortaleza, Brasil	77,34	22,66
5	Natal, Brasil	77,18	22,82
6	Salvador de Bahía, Brasil	76,65	23,35
7	Recife, Brasil	76,36	23,64
8	Puerto España, Trinidad y Tobago	76,21	23,79
9	Baltimore, Maryland, Estados Unidos	75,49	24,51
10	Rosario, Argentina	75,18	24,82
11	Memphis, Tennessee, Estados Unidos	74,79	25,21
12	Detroit, Michigan, Estados Unidos	74,12	25,88
13	Porto Alegre, Brasil	72,39	27,61
14	Guayaquil, Ecuador	71,58	28,42
15	Albuquerque, Nuevo México, Estados Unidos	71,50	28,50
16	Tijuana, México	71,44	28,56
17	Lima, Perú	70,90	29,10
18	Kingston, Jamaica	70,90	29,10
19	Saint Louis, Missouri, Estados Unidos	70,64	29,36
20	Sao Paulo, Brasil	70,49	29,51
21	Cali, Colombia	70,23	29,77
22	Ciudad de México, México	68,49	31,51

**Figura 1.1:** Índice de criminalidad de los países de América 2023

a los robos que se cometen en las calles. Si se toma en cuenta el tipo de acto delictivo que puede atentar a la vida humana en las calles, se tiene que, el robo de vehículos representa un 2 %, mientras que el secuestro y extorsión el 0.2 %; que, si bien son menos frecuentes, representan una amenaza al bienestar y seguridad de los ciudadanos de las zonas urbanas (Statista Research Department, 2024). El último informe del Organismo Supervisor de Inversión Privada en Telecomunicaciones (Osiptel) de 2023, se registraron un total de 1,706,643 casos de robo de celulares en el año 2023, el cual el mayor número de robo ha sido los días lunes (Infobae, 2024). Asimismo, se informó que, entre el periodo de enero y mayo del año 2023, por cada hora, se roban en promedio, 200 celulares a nivel nacional (El Comercio, 2023). Estos resultados no solo representan una pérdida material y económica por parte de los afectados, genera desconfianza y preocupación sobre su seguridad personal y la posible vulneración en su privacidad de información que se contenía en el dispositivo del agraviado.

Con lo que respecta al robo de vehículos, que también atenta a la vida y seguridad de la población de las zonas urbanas, según la Asociación Automotriz del Perú (APP) en el año 2023, la venta de vehículos ha representado un aumento del 2.4 % (181,000 autos vendidos) con respecto al año 2022 (AmChamnews, 2024). Este incremento en la venta de autos, también lleva consigo el aumento a los robos de vehículos, que es un problema que se está volviendo cada vez más persistente en Perú, ya que, según la Superintendencia Nacional de Registros



**Figura 1.2:** Días con más robos de celulares registrados en el 2023 Osiptel

Públicos, en promedio se roban 46 autos al día y que en el primer semestre del año 2022 se ha registrado un total de 9,758 autos robados en el Perú (Tracklink, 2023).

Estas cifras demuestran la necesidad de desarrollar herramientas y/o estrategias que permitan mitigar los riesgos relacionados con los actos delictivos en entornos rurales, donde se destaca el robo de celulares y vehículos, que, a su vez, pueden estar relacionadas entre sí. Las aplicaciones móviles para ayudar al usuario a elegir rutas más cortas para llegar a su destino en el menor tiempo posible mediante un motor de crowdsourcing (recopilación de tráfico), son cada vez más comunes y fundamentales de utilizar; ya que, en 2022, más de 150 millones de personas utilizan Waze; mientras que Google maps ha sido instalada por más de 10 mil millones de usuarios (autocosmos, 2022). Sin embargo, estas herramientas de ayuda no tienen una funcionalidad que les permita detectar entre qué calles son seguras o no tomando en cuenta información o denuncias de delitos cometidos en esas zonas, ocasionando que los usuarios queden expuestos a situaciones peligrosas como el robo de su celular, de su auto o ambas.

## 1.2. Formulación del Problema

formulación de los problemas de la presente investigación, se elaboró un «árbol de problemas» (véase Anexo A.1).

### 1.2.1. Problema General

¿Es posible utilizar un algoritmo de aprendizaje profundo para detectar rutas seguras en las calles de los Ángeles?

### 1.2.2. Problemas Específicos

- ¿Qué algoritmos de aprendizaje profundo son los más representativos de los antecedentes?
- ¿Qué alternativas de procesamiento de datos se deben aplicar a la base de datos para entrenar el modelo?
- ¿Qué métricas propuestas en los antecedentes existen para ayudar en la eficiencia y eficacia del modelo propuesto?
- ¿El modelo propuesto evita las zonas con alto nivel de delincuencia?

## 1.3. Objetivos de la Investigación

Para la formulación de los objetivos de la presente investigación se elaboró un «árbol de objetivos» (véase Anexo [A.2](#)).

### 1.3.1. Objetivo General

Desarrollar un algoritmo de aprendizaje profundo para detectar rutas seguras en las calles de los Ángeles.

### 1.3.2. Objetivos Específicos

- Seleccionar los algoritmos de aprendizaje profundo más representativos de cada antecedente para la detección de rutas seguras.
- Seleccionar una alternativa de procesamiento de datos que permita entrenar un modelo de detección de rutas seguras.
- Evaluar la factibilidad del modelo con las métricas propuestas en los antecedentes para mejorar el tiempo de procesamiento con los menores recursos.
- Demostrar que el modelo propuesto evita las zonas con alto nivel de delincuencia.

## **1.4. Justificación de la Investigación**

### **1.4.1. Teórica**

Esta investigación se realiza con el propósito de desarrollar un algoritmo que pueda realizar el procesamiento de direcciones y el cálculo de riesgo en los diferentes nodos que existen en un grafo a fin de poder generar métodos de compensación para llegar a un destino de manera segura y rápida. Además, dicho algoritmo será realizado bajo Aprendizaje profundo, para crear un modelo robusto de detección con menos tiempo de procesamiento y recursos para detectar rutas seguras que eviten zonas de alto nivel de crímenes registrados en una ciudad; asimismo, de ser un modelo escalable que no dependa de volver a pasar por el proceso de entrenamiento para cuando se quiera detectar una ruta distinta a otra ciudad.

Actualmente, existen investigaciones que utilizan modelos tradicionales como el algoritmo de Dijkstra, pero que carecen de escalabilidad y su complejidad computacional aumenta a medida que el tamaño de los datos y nuevos escenarios aumenten (Diya Lia y e.al, 2023). Por último, en el Perú, estos tipos de proyectos nos son muy comunes, por lo que el modelo propuesto puede aportar a las nuevas investigaciones en el uso de algoritmos de Aprendizaje Profundo utilizando grafos.

### **1.4.2. Práctica**

Al culminar la investigación, se podrá utilizar el algoritmo propuesto, el cuál puede recibir por parte del usuario la dirección a la cuál quiere ir y la dirección en la que se encuentra, para mostrarle en un tiempo de ejecución menor la ruta más segura a la cuál puede ir. Además, podrá tener la capacidad de identificar rutas seguras a partir nuevos datos de crímenes registrados en distintas ciudades.

La presente investigación demostrará que un modelo de aprendizaje profundo puede detectar rutas seguras en grafos cada vez más complejos y ser escalables en comparación a los métodos tradicionales, mejorando de esta manera la calidad de vida de las personas, disminuir el índice de criminalidad, informar a los usuarios que priorizan su seguridad o que no tienen conocimiento de las calles de una ciudad a la que visitan a decidir el camino que debe tomar para sentir más seguro.

### **1.4.3. Metodológica**

. La implementación de este algoritmo puede ayudar a las personas a conocer qué rutas son seguras para transportarse y evitar ser víctima de algún crimen, ya que, si bien se cuenta con aplicaciones de enrutamiento en tiempo real como Waze o Google Maps, estos no cuentan con la capacidad de detectar qué rutas son las más seguras, debido a que no trabajan con datos de registros de delitos. Asimismo, esta investigación puede aportar a la disminución del índice de criminalidad de alguna ciudad e informar a las personas qué rutas no deben tomar.

Para ello, se utilizaron técnicas de Aprendizaje Profundo aplicados en grafos que son entrenados con un conjunto de datos de crímenes registrados en las calles de una ciudad.

## **1.5. Delimitación del Estudio**

### **1.5.1. Espacial**

Para la presente investigación se realizará con los registros de actos delictivos recolectados y almacenados en la web oficial del Departamento de Policía de los Ángeles (APD) para uso público, la cual consiste en denuncias o actos registrados por personas, junto con el tipo de delincuencia y la dirección dónde ocurrió el hecho. Estos datos están registrados en formato “csv”. Además, se utilizará librerías que permitan cartografiar las calles de la ciudad de los Ángeles y poder utilizarlos como grafos.

### **1.5.2. Temporal**

La presente investigación analizará un conjunto de datos de registros delictivos de la ya mencionada base de la web oficial del APD, que empieza desde el año 2020 hasta los últimos registros descargados a la fecha que se inicia este trabajo.

### **1.5.3. Conceptual**

Esta investigación se centrará en el desarrollo de un algoritmo que logre detectar qué rutas de la ciudad de los Ángeles serán seguras para evitar algún caso de crimen basado en un indicador de riesgo. Para lograrlo, se necesitará herramientas de Aprendizaje Profundo y métodos de procesamiento de datos para desarrollar un modelo que se adapte a los diferentes estilos de grafos de forma óptima. El modelo desarrollado busca poder dar una contribución



a la comunidad con el fin de evitar situaciones peligrosas y apoyar a los distintos métodos de reducción en la tasa de criminalidad que se presenta en las zonas urbanas; asimismo aumentar la participación en técnicas de Aprendizaje Profundo utilizando grafos.

## **1.6. Hipótesis**

### **1.6.1. Hipótesis General**

Mediante el desarrollo de un algoritmo de aprendizaje profundo se logra detectar rutas seguras en las calles de los Ángeles.

### **1.6.2. Hipótesis Específicas**

- La selección de los algoritmos de aprendizaje profundo más representativos de cada antecedente mejora significativamente la precisión y eficiencia al detectar rutas seguras.
- La identificación de una alternativa de procesamiento de datos influye positivamente para poder entrenar un modelo de detección de rutas seguras.
- La utilización de métricas propuestas en los antecedentes es factible para optimizar el tiempo de procesamiento del modelo.
- Se demuestra que el modelo propuesto evita las zonas con alto nivel de delincuencia.

### **1.6.3. Matriz de Consistencia**

A continuación se presenta la matriz de consistencia elaborada para la presente investigación (véase Anexo [A.1](#)).

## Capítulo 2

# MARCO TEÓRICO

### 2.1. Antecedentes de la investigación

En esta sección se presentarán diversos artículos de investigación o tesis las cuales abordarán diversas técnicas y enfoques que se emplearon para afrontar problemas similares al de esta tesis. Asimismo, a continuación se presenta un cuadro resumen (véase Anexo [B.1](#)) de lo que se presenta en esta sección.

#### 2.1.1. Copper price estimation using bat algorithm ([Dehghani & Bogdanovic, 2018](#))

Dehghani y Bogdanovic realizaron un artículo de investigación el cual fue publicado en la revista «Resources Policy» en el año 2018. Este fue titulado «Copper price estimation using bat algorithm» la cual traducida al español significa «Estimación del precio del cobre utilizando el algoritmo bat».

##### 2.1.1.1. Planteamiento del Problema y objetivo

hhhhj

##### 2.1.1.2. Técnicas empleadas por los autores

Los autores plantearon emplear una combinación entre la función de series de tiempo y el aljhkk.

### 2.1.1.3. Metodología empleada por los autores

gfhhhh

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - T_i)^2}{N}} \quad (\text{Ecuación 2.1})$$

gfhf tal forma mejorar aún más la precisión de la predicción del precio del cobre.

### 2.1.1.4. Resultados obtenidos

Las funciones de serie de tiempo más importantes se usaron para estimar los cambios en el precio del cobre. Entre ellos, la serie BMMR con una media de RMSE de 0.449 presentó la mejor estimación. El algoritmo Bat se usó para modificar la función de tiempo BMMR debido a su alta capacidad para estimar los cambios en el precio del metal. Se obtuvo un RMSE de 0.132 de la ecuación modificada con BA. Los resultados obtenidos tienen una precisión mucho mayor y, a diferencia del BMMR, están más cerca de la realidad.

## 2.2. Bases Teóricas

### 2.2.1. Machine Learning

Es un subcampo de ejecutar dificultosos procesos aprendiendo de datos, en lugar de seguir reglas preprogramadas ([Royal Society Working Group, 2017](#)).

es importante mencionar que existen también cinco tipos de problemas de aprendizaje que se pueden enfrentar: regresión, clasificación, simulación, optimización y clusterización ([Gollapudi, 2016](#)). Por otro lado, el aprendizaje automático también posee una división por subcampos que se puede observar en la Figura 14.

### 2.2.2. Natural Language Processing (NLP)

Naturalmano ([Goyal et al., 2018](#)). Otra definición para este término implica que es un campo especializado de la informática que es

De acuerdo con [Goyal et al. \(2018\)](#), e

## **2.3. Marco Conceptual**

Para de

## Capítulo 3

# METODOLOGÍA DE LA INVESTIGACIÓN

### 3.1. Diseño de la investigación

En esta sección del documento se explicará cual es el diseño, el tipo y el enfoque del trabajo de investigación, así como también la población y la muestra.

#### 3.1.1. Diseño no experimental

El diseño es no experimental longitudinal, ya que las variables no serán manipuladas y serán analizadas tal como se encuentran. Es decir, tanto los datos textuales (noticias) y el precio del cobre serán analizados sin ningún cambio aplicando técnicas de procesamiento de lenguaje natural y algoritmos de aprendizaje automático con la finalidad de crear un modelo productivo robusto y facilitar la predicción del cobre. Asimismo, la recolección de datos que se realizará será en un determinado periodo de tiempo.

#### 3.1.2. Tipo explicativo

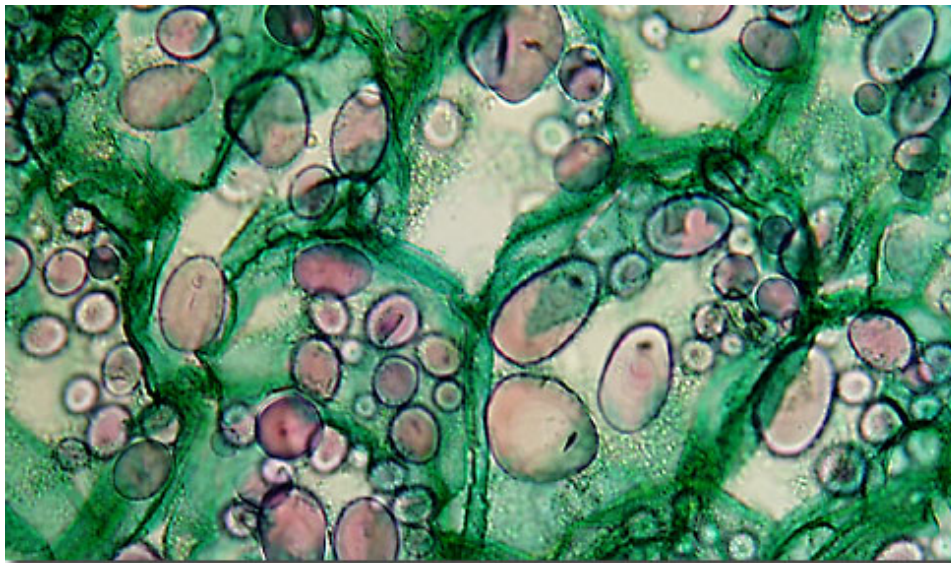
El alcance de la presente investigación es explicativo debido a que se busca explicar el comportamiento volátil del precio del cobre en base a noticias de periódicos digitales y además predecirlo.

### 3.1.3. Enfoque cuantitativo

El enfoque esta investigación es cuantitativo dado que se empleará técnicas del procesamiento de lenguaje natural (NLP), las cuales conllevan a procesar los datos de tipo textual a numéricos (vectores de características) y con ello posteriormente usar técnicas estadísticas como la regresión lineal para la predicción del precio del cobre.

## 3.2. Población y muestra

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit. La Figura 3.1 y el Cuadro 3.1



**Figura 3.1:** Prueba de Figura

## 3.3. Operacionalización de Variables

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus

mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

### 3.4. Instrumentos de medida

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat

- muscle and fat cells remove glucose from the blood,
- cells breakdown glucose via glycolysis and the citrate cycle, storing its energy in the form of ATP,
- liver and muscle store glucose as glycogen as a short-term energy reserve,
- adipose tissue stores glucose as fat for long-term energy reserve, and
- cells use glucose for protein synthesis.

### 3.5. Técnicas de recolección de datos

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

L<sup>A</sup>T<sub>E</sub>X is great at typesetting mathematics. Let  $X_1, X_2, \dots, X_n$  be a sequence of independent and identically distributed random variables with

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_i^n X_i \quad (\text{Ecuación 3.1})$$

La Ecuación [Ecuación 3.1](#) denote their mean. Then as  $n$  approaches infinity, the random variables

$$\sqrt{n}(S_n - \mu)$$

converge in distribution to a normal  $\mathcal{N}(0, \sigma^2)$ .

## 3.6. Técnicas para el procesamiento y análisis de la información

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

You can make lists with automatic numbering ...

1. Like this,
2. and like this.

... or bullet points ...

- Like this,
- and like this.

## 3.7. Cronograma de actividades y presupuesto

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

Item	Quantity
Widgets	42
Gadgets	13

**Tabla 3.1:** An example table.



# Capítulo 4

## DESARROLLO DEL EXPERIMENTO

### 4.1. X

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn? Kjift ”not at all!...

### 4.2. Y

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

Item	Quantity
Widgets	42
Gadgets	13

**Tabla 4.1:** An example table.

### 4.3. Z

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

El paper es citado y el otro paper .

## Capítulo 5

# ANÁLISIS Y DISCUSIÓN DE RESULTADOS

### 5.1. X

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn? Kjift ”not at all!...

### 5.2. Y

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

Item	Quantity
Widgets	42
Gadgets	13

**Tabla 5.1:** An example table.

### 5.3. Z

Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

## Capítulo 6

# CONCLUSIONES Y RECOMENDACIONES

### 6.1. Conclusiones

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn? Kjift ”not at all!...

### 6.2. Recomendaciones

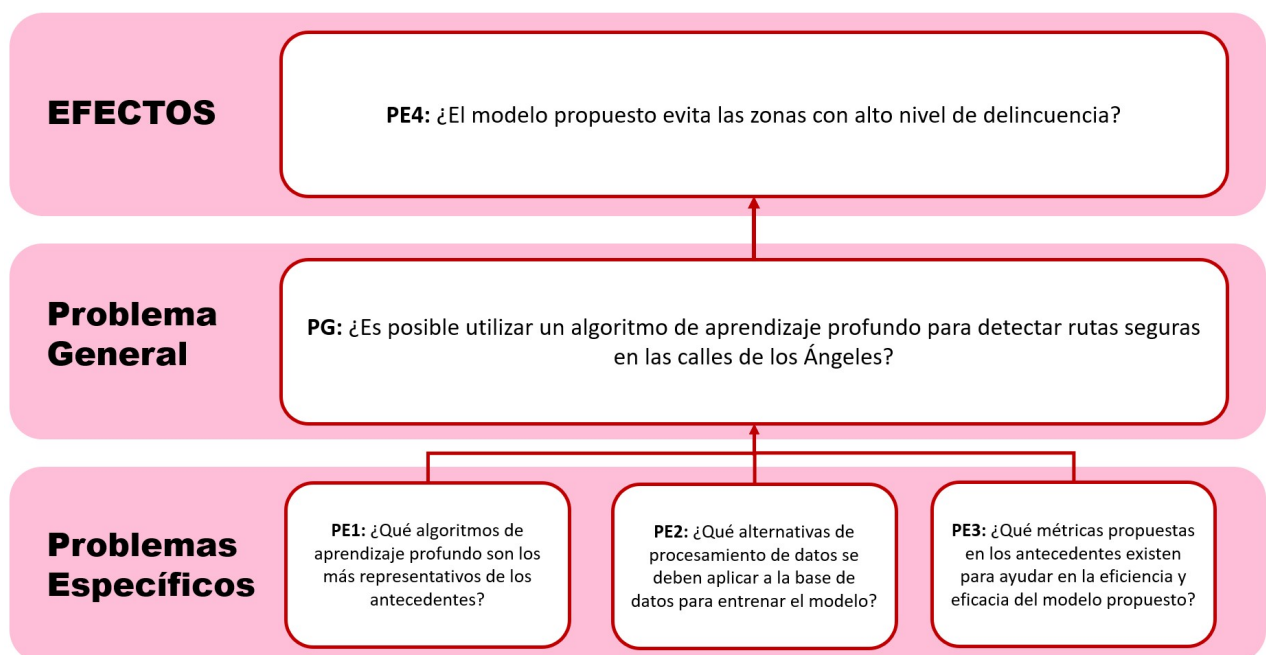
Nisi porta lorem mollis aliquam ut porttitor leo. Aenean pharetra magna ac placerat vestibulum. Est placerat in egestas erat imperdiet sed euismod. Velit euismod in pellentesque massa placerat. Enim praesent elementum facilisis leo vel fringilla. Ante in nibh mauris cursus mattis molestie a iaculis. Erat pellentesque adipiscing commodo elit at imperdiet dui accumsan sit. Porttitor lacus luctus accumsan tortor posuere ac ut. Tortor at auctor urna nunc id. A iaculis at erat pellentesque adipiscing commodo elit.

## **Anexos**

# Anexos A

## Anexo I: Matriz de Consistencia

### A.1. Árbol de Problemas

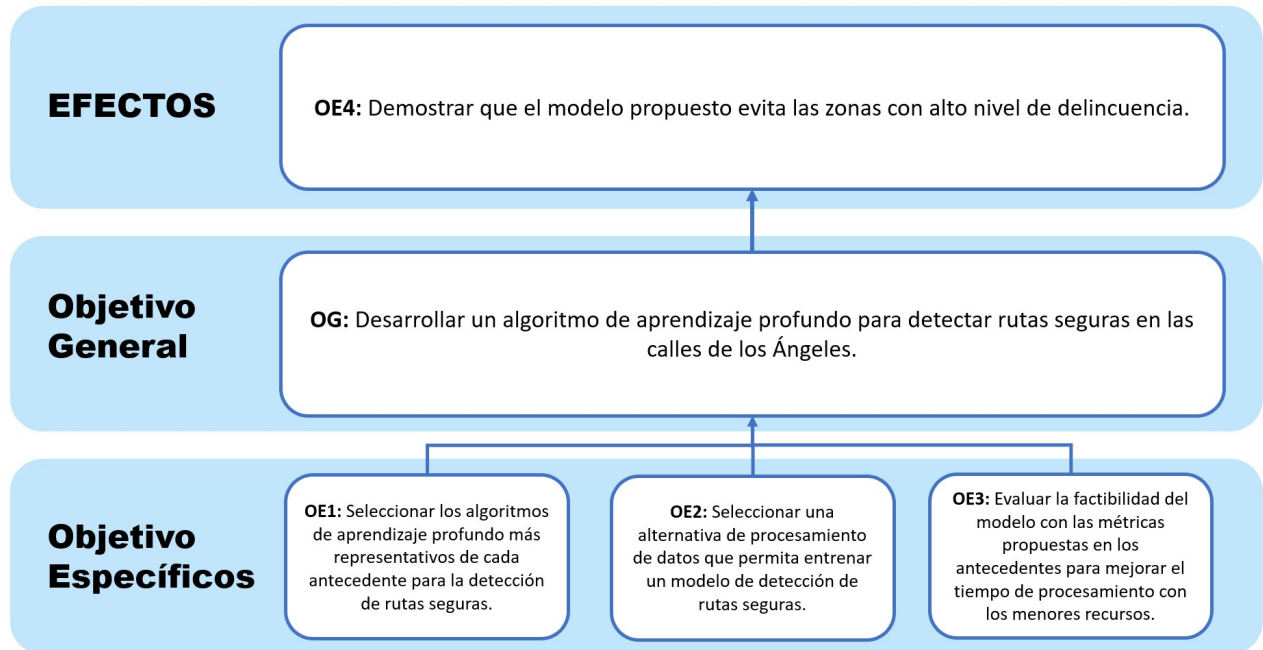


PROBLEMAS	OBJETIVOS	HIPÓTESIS
Problema General	Objetivo General	Hipótesis General
¿Es posible utilizar un algoritmo de aprendizaje profundo para detectar rutas seguras en las calles de los Ángeles?	Desarrollar un algoritmo de aprendizaje profundo para detectar rutas seguras en las calles de los Ángeles.	Mediante el desarrollo de un algoritmo de aprendizaje profundo se logra detectar rutas seguras en las calles de los Ángeles.
Problemas Específicos	Objetivos Específicos	Hipótesis Específicas
¿Qué algoritmos de aprendizaje profundo son los más representativos de los antecedentes?	Seleccionar los algoritmos de aprendizaje profundo más representativos de cada antecedente para la detección de rutas seguras.	La selección de los algoritmos de aprendizaje profundo más representativos de cada antecedente mejora significativamente la precisión y eficiencia al detectar rutas seguras.
¿Qué alternativas de procesamiento de datos se deben aplicar a la base de datos para entrenar el modelo?	Seleccionar una alternativa de procesamiento de datos que permita entrenar un modelo de detección de rutas seguras.	La identificación de una alternativa de procesamiento de datos influye positivamente para poder entrenar un modelo de detección de rutas seguras.
¿Qué métricas propuestas en los antecedentes existen para ayudar en la eficiencia y eficacia del modelo propuesto?	Evaluar la factibilidad del modelo con las métricas propuestas en los antecedentes para mejorar el tiempo de procesamiento con los menores recursos.	La utilización de métricas propuestas en los antecedentes es factible para optimizar el tiempo de procesamiento del modelo.
¿El modelo propuesto evita las zonas con alto nivel de delincuencia?	Demostrar que el modelo propuesto evita las zonas con alto nivel de delincuencia.	Se demuestra que el modelo propuesto evita las zonas con alto nivel de delincuencia.

**Tabla A.1:** Matriz de consistencia. Fuente: Elaboración propia



## A.2. Árbol de Objetivos



## **Anexos B**

### **Anexo II: Resumen de Papers investigados**

Tipo	N°	Título	Autor	Año	País	Fuente
Problema	1	Copper price estimation using bat algorithm	Dehghani Bogdanovic	2018	United Kingdom	Resources Policy
	2	Alternative techniques for forecasting mineral commodity prices	Cortez, Saydam, Coulton, Sammut	2018	Netherlands	International Journal of Mining Science and Technology
Propuesta	3	Prediction of the crude oil price thanks to natural language processing applied to newspapers	Trastour, Genin, Morlot	2016	USA	Standfort University ML repository
	4	Stock Price Prediction Using Deep Learning	Tipirisetty	2018	USA	Master's Theses San Jose State University
	5	Deep Learning for Stock Prediction Using Numerical and Textual Information	Akita, R., Yoshihara, A., Matsubara, T., Uehara, K.	2016	USA	2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)
Técnica	6	Stock Prices Prediction using the Title of Newspaper Articles with Korean Natural Language Processing	Yun, Sim, Seok	2019	Japan	2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)
	7	A Method of Optimizing LDA Result Purity Based on Semantic Similarity	Jingrui, Z., Qinglin, W., Yu, L., Yuan, L.	2017	China	2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)
	8	Qualitative Stock Market Predicting with Common Knowledge Based Nature Language Processing: A Unified View and Procedure	Rao, D., Deng, F., Jiang, Z., Zhao, G.	2015	USA	2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics
	9	Fuzzy Bag-of-Words Model for Document Representation	Zhao, R., Mao, K.	2018	USA	IEEE Transactions on Fuzzy Systems ( Volume: 26 , Issue: 2 , April 2018 )

**Tabla B.1:** Cuadro Resumen de Papers investigados. Fuente: Elaboración propia

# BIBLIOGRAFÍA

- Akita, R., Yoshihara, A., Matsubara, T., & Uehara, K. Deep learning for stock prediction using numerical and textual information. En: *En 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*. IEEE. 2016, 1-6.
- Cortez, C. T., Saydam, S., Coulton, J., & Sammut, C. (2018). Alternative techniques for forecasting mineral commodity prices. *International Journal of Mining Science and Technology*, 28(2), 309-322.
- Dehghani, H., & Bogdanovic, D. (2018). Copper price estimation using bat algorithm. *Resources Policy*, 55, 55-61.
- Gartner. (2019). Gartner IT Glossary. <https://www.gartner.com/it-glossary/>
- Gollapudi, S. (2016). *Practical machine learning*. Packt Publishing Ltd.
- Goyal, P., Pandey, S., & Jain, K. (2018). Deep learning for natural language processing. *Deep Learning for Natural Language Processing: Creating Neural Networks with Python [Berkeley, CA]: Apress*, 138-143.
- IBM. (2019). IBM AI glossary. <https://www.ibm.com/cloud/garage/architectures/cognitiveArchitecture/glossary>
- Jingrui, Z., Qinglin, W., Yu, L., & Yuan, L. A method of optimizing LDA result purity based on semantic similarity. En: *En 2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE. 2017, 361-365.
- Kulkarni, A., & Shivananda, A. (2019). Exploring and Processing Text Data. En *Natural Language Processing Recipes* (pp. 37-65). Springer.
- Lagos, G. (2017). ¿Cómo lo han hecho los especialistas?: aciertos y desaciertos al proyectar el precio del cobre. <https://gyn.claseejecutiva.uc.cl/como-lo-han-hecho-los-especialistas-aciertos-y-desaciertos-al-proyectar-el-precio-del-cobre/#>
- Martínez, R., & Cohen, E. (2018). Manual formulación, evaluación y monitoreo de proyectos sociales. <https://dds.cepal.org/redesoc/publicacion?id=242>
- Ministerio de Energía y Minas (ESTAMIN). (2019). *Perú: país líder de los metales del futuro* (Boletín Estadístico Minero).

- Rao, D., Deng, F., Jiang, Z., & Zhao, G. Qualitative Stock Market Predicting with Common Knowledge Based Nature Language Processing: A Unified View and Procedure. En: En *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*. 2. IEEE. 2015, 381-384.
- Real Academia Española. (2014). Diccionario de la lengua española (Twenty-third). <https://dle.rae.es/?w=diccionario>
- Royal Society Working Group. (2017). *Machine learning: the power and promise of computers that learn by example* (inf. téc.). Technical report.
- Study Group International Copper. (2018). The World Copper Factbook 2018. [www.icsg.org](http://www.icsg.org)
- TensorFlow. (2019). Vector Representations of Words. <https://www.tensorflow.org/tutorials/representation/word2vec>
- Tipirisetty, A. (2018). *Stock Price Prediction using Deep Learning* [Tesis de maestría, San José State University] [Master's Projects]. <https://doi.org/https://doi.org/10.31979/etd.bzmm-36m7>
- Trastour, S., Genin, M., & Morlot, A. (2016). Prediction of the crude oil price thanks to natural language processing applied to newspapers. <http://cs229.stanford.edu/proj2016/report/>
- U.S. Geological Survey. (2019). Mineral commodity summaries 2019: U.S. Geological Survey. <https://doi.org/https://doi.org/10.3133/70202434>
- Yun, H., Sim, G., & Seok, J. Stock Prices Prediction using the Title of Newspaper Articles with Korean Natural Language Processing. En: En *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*. IEEE. 2019, 019-021.
- Zhao, R., & Mao, K. (2017). Fuzzy bag-of-words model for document representation. *IEEE Transactions on Fuzzy Systems*, 26(2), 794-804.