



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

Detección de audio deepfake en Español utilizando redes neuronales profundas: Análisis de tono, timbre y patrones de voz

Trabajo de investigación para el curso de Trabajo de Tesis I

Antonio Jesús Barrera Benetres
Asesor: Marks Calderón

Lima, 30 de septiembre 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. Faciliis 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. 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.

Índice general

Índice de Figuras	8
Índice de Tablas	9
1. PLANTEAMIENTO DEL PROBLEMA	10
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	12
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	13
1.4.1. Teórica	13
1.4.2. Práctica	14
1.4.3. Metodológica	14
1.5. Delimitación del Estudio	15
1.5.1. Espacial	15
1.5.2. Temporal	16
1.5.3. Conceptual	16

1.6. Hipótesis	16
1.6.1. Hipótesis General	16
1.6.2. Hipótesis Específicas	16
1.6.3. Matriz de Consistencia	17
2. MARCO TEÓRICO	18
2.1. Antecedentes de la investigación	18
2.1.1. Copper price estimation using bat algorithm (pr`dehghani2018copper)	18
2.2. Bases Teóricas	19
2.2.1. Machine Learning	19
2.2.2. Natural Language Processing (NLP)	19
2.3. Marco Conceptual	20
3. METODOLOGÍA DE LA INVESTIGACIÓN	21
3.1. Diseño de la investigación	21
3.1.1. Diseño no experimental	21
3.1.2. Tipo explicativo	21
3.1.3. Enfoque cuantitativo	22
3.2. Población y muestra	22
3.3. Operacionalización de Variables	22
3.4. Instrumentos de medida	23
3.5. Técnicas de recolección de datos	23
3.6. Técnicas para el procesamiento y análisis de la información	24
3.7. Cronograma de actividades y presupuesto	24
4. DESARROLLO DEL EXPERIMENTO	25
4.1. X	25

4.2. Y	25
4.3. Z	26
5. ANÁLISIS Y DISCUSIÓN DE RESULTADOS	27
5.1. X	27
5.2. Y	27
5.3. Z	28
6. CONCLUSIONES Y RECOMENDACIONES	29
6.1. Conclusiones	29
6.2. Recomendaciones	29
Anexos	30
A. Anexo I: Matriz de Consistencia	31
B. Anexo II: Resumen de Papers investigados	33

Índice de Figuras

3.1. Prueba de Figura	22
---------------------------------	----

Índice de Tablas

3.1. An example table.	24
4.1. An example table.	25
5.1. An example table.	27
A.1. Matriz de consistencia. Fuente: Elaboración propia	32
B.1. Cuadro Resumen de Papers investigados. Fuente: Elaboración propia	34

Capítulo 1

PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción de la Realidad Problemática

En la actualidad, el uso de tecnologías avanzadas, como la inteligencia artificial (IA), ha permitido la creación de contenido multimedia falsificado conocido como deepfake. Este contenido incluye videos, imágenes y, más recientemente, audios. Los deepfakes de audio han adquirido una relevancia particular debido a su capacidad para imitar voces humanas con una precisión asombrosa, lo que ha desencadenado una serie de problemas relacionados con la seguridad y la confianza en las comunicaciones digitales (**heidari2023**). Esta tecnología ha comenzado a ser utilizada para fraudes y suplantación de identidad, presentando un desafío significativo para la detección de audios manipulados, especialmente en español, donde las herramientas actuales de detección no están adecuadamente adaptadas para capturar las particularidades del idioma.

En Perú, el fraude mediante deepfakes de audio ha mostrado un aumento considerable en los últimos años. Un informe de la Policía Nacional del Perú (PNP) revela que más de un millón de soles han sido robados mediante la clonación de voces utilizando IA. En estos fraudes, los delincuentes suelen imitar la voz de familiares o amigos para solicitar dinero en situaciones de emergencia falsas, lo que resulta en un engaño muy difícil de detectar por las víctimas (**rojas2023**). Este incremento de fraudes plantea un problema general relacionado con la incapacidad de las herramientas actuales para detectar audios falsificados, lo que facilita la comisión de delitos en el ámbito digital.

Uno de los problemas específicos más críticos es la falta de un dataset en español que contemple variaciones regionales y voces manipuladas para entrenar modelos de redes neuronales profundas. La mayoría de los avances en la detección de deepfakes de audio se han

centrado en el inglés, lo que genera una brecha importante en la capacidad de los modelos para adaptarse a las características del habla en español, que presenta variaciones significativas en términos de patrones de voz, frecuencia fundamental y ritmo del habla (**heidari2023**). Esta falta de datos específicos en español dificulta la construcción de modelos robustos que puedan detectar deepfakes con precisión en este idioma.

Otro problema específico está relacionado con las limitaciones técnicas de las herramientas actuales para analizar variables clave como el tono de voz, timbre de voz y formantes. Los audios falsificados mediante deepfakes logran imitar estas características acústicas de manera casi perfecta, lo que confunde tanto a los sistemas de detección como a las personas. En particular, los deepfakes en español plantean desafíos únicos debido a las diferencias fonéticas con otros idiomas, lo que complica aún más la detección efectiva de manipulaciones (**amesquita2023**).

Un tercer problema específico se refiere al uso de deepfakes de audio en fraudes por suplantación de identidad en Perú, donde los delincuentes utilizan esta tecnología para hacer pasar sus voces por las de familiares o colegas en situaciones de emergencia o negociación. Estos fraudes, que afectan tanto a individuos como a empresas, se ven agravados por la dificultad de las técnicas actuales para analizar características como la prosodia, articulación, transiciones entre fonemas y ruidos de fondo. Estos factores son críticos para la identificación precisa de audios falsificados, ya que los deepfakes no siempre logran replicar con fidelidad estos aspectos del habla humana, pero los sistemas de detección existentes no están optimizados para capturarlos (**chen2024**).

A nivel global, se ha documentado un incremento en el uso de deepfakes para cometer fraudes, con pérdidas millonarias en varios países. En Hong Kong, por ejemplo, se reportó un caso en el que un trabajador fue engañado mediante una videollamada con deepfakes de varios miembros de la junta directiva de su empresa, lo que resultó en un fraude de 25 millones de dólares (**chen2024**). Este tipo de casos destaca la urgencia de desarrollar soluciones tecnológicas más efectivas que permitan la detección de deepfakes de audio en español, donde las herramientas actuales siguen siendo insuficientes.

En el contexto peruano, la proliferación de fraudes basados en deepfakes de audio no solo afecta a los individuos, sino también a figuras públicas. Un caso reciente involucró a la presidenta Dina Boluarte, cuya voz fue manipulada mediante IA para hacer parecer que promovía una inversión fraudulenta, lo que generó confusión entre el público y demostró el poder de esta tecnología para influir en la opinión pública y causar daño reputacional (**amesquita2023**).

Por todo lo anterior, es evidente que la detección de deepfakes de audio en español requiere un enfoque más sofisticado. El desarrollo de modelos basados en redes neuronales pro-

fundas que puedan analizar variables acústicas clave, como el tono de voz, timbre de voz, patrones de voz, frecuencia fundamental, duración y ritmo del habla, formantes, nivel de energía del habla, ruidos de fondo, prosodia, articulación y transiciones entre fonemas, es esencial para mitigar el impacto de los fraudes por deepfakes en Perú. Estas variables juegan un papel fundamental en la autenticidad del habla y pueden proporcionar pistas valiosas para identificar audios manipulados. Sin embargo, la falta de modelos especializados en español y la escasez de datasets específicos siguen siendo los principales obstáculos para lograr una detección precisa y confiable.

1.2. Formulación del Problema

1.2.1. Problema General

El incremento del fraude en Perú mediante el uso de tecnologías deepfake de audio ha evidenciado la falta de herramientas adecuadas para detectar estos fraudes, especialmente en español. Las técnicas actuales no logran identificar eficazmente las características acústicas del español, como el tono de voz, timbre de voz, patrones de voz, frecuencia fundamental (pitch), duración y ritmo del habla, formantes, nivel de energía del habla (intensidad), ruidos de fondo, prosodia, articulación y transiciones entre fonemas, lo que facilita la suplantación de identidad y el fraude en las comunicaciones personales y empresariales.

1.2.2. Problemas Específicos

- La falta de un dataset en español que incluya variaciones regionales y voces manipuladas dificulta el entrenamiento de modelos de redes neuronales profundas para detectar deepfakes de audio en español, debido a las diferencias en patrones de voz, frecuencia fundamental (pitch) y ritmo del habla.
- Las técnicas actuales no logran detectar las variaciones en el tono de voz, timbre de voz y formantes en audios en español, lo que disminuye la precisión en la identificación de audios manipulados.
- Los fraudes por suplantación de identidad mediante deepfakes de audio en Perú son difíciles de detectar con las técnicas actuales debido a la falta de análisis de prosodia, articulación, transiciones entre fonemas y ruidos de fondo, lo que incrementa el riesgo de fraude.

1.3. Objetivos de la Investigación

Para la formulación de los objetivos de la presente investigación se elaboró un «árbol de objetivos» (véase Anexo 2)

1.3.1. Objetivo General

Desarrollar un modelo basado en redes neuronales profundas que permita detectar deepfakes de audio en español mediante el análisis de variables clave como tono de voz, timbre de voz, patrones de voz, frecuencia fundamental, duración y ritmo del habla, formantes, nivel de energía del habla, ruidos de fondo, prosodia, articulación y transiciones entre fonemas, mejorando la precisión en la identificación de audios manipulados para mitigar fraudes por suplantación de identidad en Perú.

1.3.2. Objetivos Específicos

- Desarrollar un dataset específico en español, con variaciones regionales y voces manipuladas, para entrenar un modelo de redes neuronales profundas que detecte deepfakes de audio
- Implementar un modelo de redes neuronales profundas que analice el tono de voz, timbre de voz y formantes para mejorar la precisión en la detección de deepfakes de audio en español.
- Evaluar la eficacia del modelo de redes neuronales profundas en la detección de deepfakes de audio en contextos de fraude por suplantación de identidad en Perú, considerando prosodia, articulación, transiciones entre fonemas y ruidos de fondo.

1.4. Justificación de la Investigación

1.4.1. Teórica

Desde una perspectiva teórica, esta investigación aporta al campo de la inteligencia artificial y el procesamiento de señales de audio, profundizando en el análisis y la detección de deepfakes mediante el uso de redes neuronales profundas. Actualmente, la mayoría de los avances en detección de deepfakes de audio se han centrado en el inglés, lo que deja una

brecha significativa en el desarrollo de modelos efectivos para el idioma español (**heidari2023**). Esta investigación se propone cubrir esa brecha, explorando cómo las características fonéticas del español (como la frecuencia fundamental, prosodia, formantes, y ruidos de fondo) pueden ser integradas en modelos de detección de audio. El desarrollo teórico de esta investigación se basa en los principios del aprendizaje profundo (deep learning) y su aplicación al análisis acústico, ampliando las bases teóricas sobre cómo los modelos de redes neuronales pueden ser adaptados para trabajar con diferentes idiomas y contextos culturales. Además, al enfocarse en las características específicas del español, esta investigación contribuye al desarrollo de datasets y herramientas especializadas que pueden ser utilizados en investigaciones futuras, tanto en el ámbito académico como en la industria

1.4.2. Práctica

La investigación sobre la detección de audios deepfake en español utilizando redes neuronales profundas es esencial en el contexto actual de seguridad digital, especialmente en Perú, donde se ha registrado un incremento preocupante de fraudes utilizando esta tecnología. Las suplantaciones de identidad mediante la clonación de voces, como lo demuestran los más de 94 casos de estafa reportados en 2023, con pérdidas superiores a un millón de soles (**rojas2023**), subrayan la vulnerabilidad de los sistemas actuales de detección. A nivel práctico, esta investigación tiene un impacto directo en la mitigación de estos fraudes. Al desarrollar modelos adaptados a las características específicas del español, como el tono de voz, timbre de voz y patrones de voz, se busca ofrecer soluciones tecnológicas robustas que permitan la identificación temprana de audios manipulados, reduciendo así los riesgos financieros y personales para los usuarios. Además, instituciones gubernamentales y empresas peruanas podrán aplicar los resultados de esta investigación para mejorar sus sistemas de seguridad y protección contra la suplantación de identidad, brindando un entorno más seguro para las comunicaciones digitales

1.4.3. Metodológica

La elección de redes neuronales profundas como metodología principal para la detección de audios deepfake se fundamenta en su capacidad superior para procesar y analizar grandes volúmenes de datos complejos, como las características acústicas de la voz humana. Las redes neuronales profundas permiten extraer automáticamente patrones y características de los datos de audio, como el tono de voz, timbre de voz, patrones de voz, frecuencia fundamental, prosodia, entre otros, lo que las convierte en una herramienta eficaz para abordar la naturaleza compleja de los deepfakes (**heidari2023**). Además, estas redes tienen la capacidad de aprender

de manera no supervisada y de adaptarse a las particularidades del idioma español, lo que es crucial dado que la mayoría de los avances en detección de deepfakes se han desarrollado en inglés.

Metodológicamente, el enfoque basado en redes neuronales profundas es ideal porque permite trabajar con un conjunto de variables independientes que incluyen tanto aspectos del habla (como la articulación y las transiciones entre fonemas) como la calidad del entorno de grabación (como los ruidos de fondo). La capacidad de estas redes para analizar simultáneamente múltiples dimensiones del audio y encontrar patrones que escapan a la detección humana o métodos tradicionales es fundamental para enfrentar la creciente sofisticación de los fraudes con deepfakes de audio (ji2024).

Otra razón metodológica clave es que el uso de redes neuronales profundas facilita la construcción de modelos predictivos que mejoran su precisión con el tiempo a medida que se entrenan con más datos. Esta flexibilidad es esencial, ya que el desarrollo de datasets específicos en español, con variaciones regionales y diferentes características de la voz, permitirá entrenar a los modelos de manera más eficaz para identificar manipulaciones en contextos reales (rojas2023). Por lo tanto, la metodología propuesta no solo es adecuada para la detección precisa de deepfakes de audio, sino que también es escalable y adaptable a futuros avances en la tecnología de falsificación de audios.

1.5. Delimitación del Estudio

1.5.1. Espacial

La presente investigación se desarrollará principalmente en el contexto de Perú, donde se ha identificado un incremento significativo en los fraudes utilizando tecnologías de deepfake de audio. El enfoque estará en analizar audios en español y las variaciones regionales del idioma dentro del país, ya que las características fonéticas del español peruano presentan particularidades que deben ser consideradas al desarrollar modelos de detección de deepfakes. Además, la aplicación de los resultados será relevante para instituciones gubernamentales, empresas y usuarios individuales en Perú, que enfrentan amenazas crecientes en el ámbito de la seguridad digital.

1.5.2. Temporal

El estudio abarcará el periodo 2023-2024, un lapso en el cual se han observado incrementos importantes en el uso fraudulento de deepfakes de audio tanto a nivel global como en Perú. La recolección de datos, construcción del dataset y desarrollo del modelo de redes neuronales profundas se llevarán a cabo en este periodo. Asimismo, se buscará analizar los fraudes recientes, reportados en los últimos dos años, para entender los patrones y tendencias de los ataques con deepfakes en audio

1.5.3. Conceptual

La investigación se centra en la detección de deepfakes de audio mediante el uso de redes neuronales profundas. Los conceptos clave abordados incluyen las características acústicas del habla, como el tono de voz, timbre de voz, frecuencia fundamental (pitch), duración y ritmo del habla, prosodia, formantes, articulación y ruidos de fondo, entre otros. Además, se explorará cómo estas variables son manipuladas por tecnologías de deepfake para suplantar identidades. El estudio también se delimita a la detección en español, debido a las diferencias fonéticas y lingüísticas entre este idioma y otros en los que se ha enfocado la investigación previa, como el inglés.

1.6. Hipótesis

1.6.1. Hipótesis General

El uso de un modelo basado en redes neuronales profundas que analice las variables acústicas clave como tono de voz, timbre de voz, patrones de voz, frecuencia fundamental, duración y ritmo del habla, formantes, nivel de energía del habla, ruidos de fondo, prosodia, articulación y transiciones entre fonemas mejora significativamente la precisión en la detección de deepfakes de audio en español, reduciendo el riesgo de fraudes por suplantación de identidad en Perú.

1.6.2. Hipótesis Específicas

- La creación de un dataset en español que incluya variaciones regionales y voces manipuladas mejorará significativamente la capacidad de las redes neuronales profundas para

detectar deepfakes de audio en este idioma

- El análisis del tono de voz, timbre de voz y formantes mediante redes neuronales profundas aumentará la precisión en la detección de deepfakes de audio en español.
- El modelo de redes neuronales profundas será más efectivo en la detección de deepfakes en contextos de fraude en Perú al incluir el análisis de prosodia, articulación, transiciones entre fonemas y ruidos de fondo, en comparación con las técnicas actuales.

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 (pr`dehghani2018copper)

pr`dehghani2018copper realizaron un artículo de investigación el cual fue publicado en la revista «Resources Policy» en el año 2018. Este fue titulado **pr`dehghani2018copper** 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 (tec'royal2017machine).

es importante mencionar que existen también cinco tipos de problemas de aprendizaje que se pueden enfrentar: regresión, clasificación, simulación, optimización y clusterización (bk'gollapudi2016practical). 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 (bk'goyal2018deep). Otra definición para este término implica que es un campo especializado de la informática que es

De acuerdo con bk'goyal2018deep<empty citation>, 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

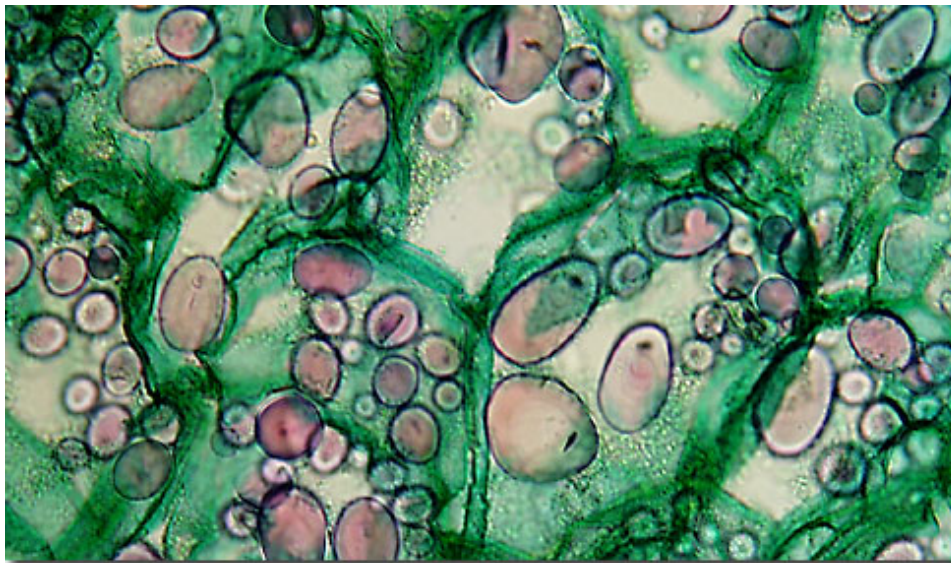


Figura 3.1: Prueba de Figura

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^AT_EX 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

PROBLEMAS	OBJETIVOS	HIPÓTESIS
Problema General	Objetivo General	Hipótesis General
El incremento del fraude en Perú mediante el uso de tecnologías deepfake de audio ha evidenciado la falta de herramientas adecuadas para detectar estos fraudes, especialmente en español. Las técnicas actuales no logran identificar eficazmente las características acústicas del español, como el tono de voz, timbre de voz, patrones de voz, frecuencia fundamental (pitch), duración y ritmo del habla, formantes, nivel de energía del habla (intensidad), ruidos de fondo, prosodia, articulación y transiciones entre fonemas, lo que facilita la suplantación de identidad y el fraude en las comunicaciones personales y empresariales.	Desarrollar un modelo basado en redes neuronales profundas que permita detectar deepfakes de audio en español mediante el análisis de variables clave como tono de voz, timbre de voz, patrones de voz, frecuencia fundamental, duración y ritmo del habla, formantes, nivel de energía del habla, ruidos de fondo, prosodia, articulación y transiciones entre fonemas, mejorando la precisión en la identificación de audios manipulados para mitigar fraudes por suplantación de identidad en Perú.	El uso de un modelo basado en redes neuronales profundas que analice las variables acústicas clave como tono de voz, timbre de voz, patrones de voz, frecuencia fundamental, duración y ritmo del habla, formantes, nivel de energía del habla, ruidos de fondo, prosodia, articulación y transiciones entre fonemas mejora significativamente la precisión en la detección de deepfakes de audio en español, reduciendo el riesgo de fraudes por suplantación de identidad en Perú.
Problemas Específicos	Objetivos Específicos	Hipótesis Específicas
La falta de un dataset en español que incluya variaciones regionales y voces manipuladas dificulta el entrenamiento de modelos de redes neuronales profundas para detectar deepfakes de audio en español, debido a las diferencias en patrones de voz, frecuencia fundamental (pitch) y ritmo del habla.	Desarrollar un dataset específico en español, con variaciones regionales y voces manipuladas, para entrenar un modelo de redes neuronales profundas que detecte deepfakes de audio	La creación de un dataset en español que incluya variaciones regionales y voces manipuladas mejorará significativamente la capacidad de las redes neuronales profundas para detectar deepfakes de audio en este idioma
Las técnicas actuales no logran detectar las variaciones en el tono de voz, timbre de voz y formantes en audios en español, lo que disminuye la precisión en la identificación de audios manipulados.	Implementar un modelo de redes neuronales profundas que analice el tono de voz, timbre de voz y formantes para mejorar la precisión en la detección de deepfakes de audio en español.	El análisis del tono de voz, timbre de voz y formantes mediante redes neuronales profundas aumentará la precisión en la detección de deepfakes de audio en español.
Los fraudes por suplantación de identidad mediante deepfakes de audio en Perú son difíciles de detectar con las técnicas actuales.	Evaluar la eficacia del modelo de redes neuronales profundas en la detección de deepfakes de audio	El modelo de redes neuronales profundas será más efectivo en la detección de deepfakes en contextos de fraude en

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