



Universidad
Tecnológica
del Perú

Lima Centro

Ingeniería de Software

Proyecto de Trabajo de Investigación

*Desarrollo de un algoritmo de inteligencia artificial para la detección de noticias falsas en redes
sociales*

Integrantes

Estrada Flores, Erick Alexander (U18101198)

Docentes

Zuñiga Vargas, Cristofher (Docente Metodólogo)

Vera Cuya, Ronald Martin (Docente de Especialidad)

Lima, Perú

Ciclo
2021-II

ÍNDICE

1. PROBLEMA DE INVESTIGACIÓN	3
2. OBJETIVO GENERAL Y ESPECÍFICOS.....	3
3. JUSTIFICACIÓN.....	3
4. REVISIÓN DE LA LITERATURA ACTUAL O ESTADO DEL ARTE	3
5. MARCO TEÓRICO	6
6. HIPÓTESIS	9
7. METODOLOGÍA	9
8. CRONOGRAMA DE TRABAJO	9
9. PRESUPUESTO.....	9
10. BIBLIOGRAFÍA.....	9
11. ANEXOS.....	11

1. PROBLEMA DE INVESTIGACIÓN

En la actualidad, el uso de las redes sociales como Facebook, Tik Tok, Instagram y Twitter, la cantidad de usuarios crece constantemente a diario. Asimismo, se está evidenciando cada vez más, la propagación de noticias falsas a través de las redes sociales, en su mayoría por motivos personales o políticos, impactando negativamente en las personas (Zervopoulos et al. 2021). Sin embargo, en los últimos años se ha evidenciado diversas investigaciones de detección de noticias falsas en redes sociales basado en técnicas de aprendizaje automático, destacando entre ellas la regresión logística como una de las más efectivas, y así poder reducir el impacto social en las personas (Desouky Fattoh and Mousa 2021). Por este motivo, en la siguiente investigación surge la siguiente pregunta: ¿De qué manera la metodología de regresión logística basado en el aprendizaje automático podría predecir las noticias falsas en redes sociales durante los dos últimos años?

2. OBJETIVO GENERAL Y ESPECÍFICOS

2.1. **Objetivo general**

2.2. **Objetivos específicos**

3. JUSTIFICACIÓN

4. REVISIÓN DE LA LITERATURA ACTUAL O ESTADO DEL ARTE

En la actualidad existen diferentes canales disponibles de redes sociales. La difusión de noticias falsas se ha vuelto cada vez más constante en las redes sociales debido a la alta disponibilidad y poca restricción de este. Por un lado, se está generando graves efectos negativos en la sociedad. Además, esto está provocando confusión en las personas, difamaciones y amenazas potenciales. Por otro lado, debido al impacto que esto está causando en la sociedad, proponen la elaboración de un modelo basado en inteligencia artificial para la detección de noticias falsas en redes sociales y reducir su impacto en la sociedad (Mokhtar et al. 2019).

Por un lado, Mahabub (2020) en su investigación propone un método para la detección de noticias falsas en redes sociales, las cuales utilizaron once algoritmos de aprendizaje automático, tales como bosque aleatorio, regresión logística, Artificial Neural Network, entre otros. Por un lado, iniciaron con un conjunto de datos que fueron recopilados con el tiempo, las cuales 3259 fueron noticias reales y 3252 noticias falsas, lo cual utilizaron para realizar dicha investigación. Asimismo, se utilizaron para la comprobación y la elaboración del estudio utilizaron once técnicas de algoritmos de clasificación de aprendizaje automático, las cuales se utilizaron en la validación

cruzada. Por otro lado, luego de que midieron la puntuación de la validación cruzada con once clasificadores, el modelo de regresión logística logró obtener el mayor puntaje con un 91.97% de puntuación, luego eligieron los tres mejores algoritmos de aprendizaje automático, los cuales fueron Regresión Logística, perceptrón Multicapa, X-Gradient Boosting. Por el lado de la clasificación de regresión logística, ajustando parámetros en función de la iteración máxima, penalización, tolerancia, solucionador y escala de intercepción, la cual obtuvo la mejor precisión con un puntaje del 93.03%. De la misma forma, Mokhtar et al. (2019) en su investigación, evidenciaron que hubo una gran diferencia entre el enfoque basado en la postura y el contenido, de las cuales la postura obtuvo un mayor rendimiento, por lo que decidieron basar el enfoque del modelo de regresión logística en la postura. Por un lado, se dieron cuenta que el modelo no se ajustaba debido a que el entrenamiento del algoritmo se realizó con un reducido número de datos, la cual no pudo predecir correctamente. Por otro lado, demuestran que el modelo de regresión logística es un excelente modelo basado en el entrenamiento de información pasada para la predicción de noticias futuras.

Asimismo, Desouky y Mousa (2021) proponen un modelo de clasificación que pueda detectar noticias falsas basado en la incrustación de Word2vec y Doc2vec como métodos de extracción de características, para lo cual ejecutaron los algoritmos mencionados en el conjunto de datos. Por un lado, evidenciaron que la incrustación de Doc2vec resultó tener una muy buena precisión en todos los algoritmos excepto LSTM, mientras que la incrustación de Word2vec resultó tener una buena precisión solo con el clasificador LSTM. Asimismo, luego de las pruebas de precisión dieron como resultado para SVM un 59,5% y para la regresión logística un 94,7%; además aplicaron Word2vec y ordenaron la palabra más común que se encontraron en el conjunto de datos, lo que resultó que era LSTM era la más adecuada, la cual obtuvo una precisión de 94.3%. Por otro lado, respecto a la precisión de los clasificadores basados en Doc2vec superan a los clasificadores basados en TF-IDF. Asimismo, el resultado de máquina de vectores de soporte obtuvo una mejor precisión con un 95,5%, luego la regresión logística con un porcentaje del 94,7% y la memoria a largo y corto plazo obtuvo una precisión más baja que las demás. Del mismo modo, Hansrajh, Adeliyi y Wing (2021) proponen un modelo de aprendizaje automático basado en la regresión logística, análisis discriminante lineal, regresión de cresta, máquina de vectores y descenso de gradiente estocástico. Por un lado, los investigadores tomaron el conjunto de datos ISOT y Liar. Asimismo, se utilizaron seis medidas de rendimiento como AUC, puntaje F1, ROC, precisión, exactitud y recuperación. Por otro lado, el modelo con mejor rendimiento respecto al conjunto de datos Liar fue el clasificador de regresión logística, que logró los mejores resultados en seis métricas de comparación, las cuales fueron AUC, ROC, exactitud y precisión, mientras que el modelo con mejor rendimiento respecto al conjunto de datos ISOT

en seis métricas de comparación como AUC, ROC, recuperación, precisión y Puntaje F1 fue el clasificador de máquina de vectores de soporte lineal.

Por un lado, Mahlous y Al-Laith (2021) proponen un modelo de aprendizaje automático capaz de detectar noticias falsas en redes sociales como Twitter, logrando ayudar a los gobiernos y responsables de la toma de decisiones del juzgado público, los investigadores se basaron en más de siete millones de tweets árabes que guardan relación con la pandemia del COVID-19. Además, usaron dos verificadores de hechos, tales como la Autoridad Antirumores de Arabia Saudita y la Agencia de Prensa de Francia para la extracción de palabras claves que puedan estar relacionadas con las noticias falsas. Por un lado, utilizaron seis clasificadores de aprendizaje automático como Regresión Logística, Naive Bayes, Perceptrón Multicapa, Máquinas de vectores de soporte, XGB y **boque** aleatorio, las cuales como resultado evidenciaron que el clasificador de regresión logística pudo obtener el mejor rendimiento de clasificación con un porcentaje del 87.8% para la clasificación de puntuación F1 y cuando se aplicó al corpus anotado automáticamente, logró un puntaje del 93,3%. Igualmente, Ali et al. (2021) realizan una investigación sobre una solución que pueda detectar noticias falsas basados en aprendizaje automático. Por un lado, **basado** en cuatro modelos **basados** en perceptrón multicapa, red neuronal recurrente, red neuronal convolucional y una híbrida entre Red neuronal convolucional y red neuronal recurrente, las cuales fueron entregadas bajo ataques como Text Fooler, Text Bugger, Deep World Bug y PWWS. Por otro lado, respecto a la precisión del modelo perceptrón multicapa, red neuronal recurrente, red neuronal convolucional y la híbrida CNN-RNN, se evidencia que perceptrón multicapa tuvo un bajo rendimiento ante diferentes ataques. Asimismo, la red neuronal convolucional y la red neuronal recurrente fueron las más robustas ante diferentes ataques basados en el conjunto de datos Liar. Por otro lado, Jiang et al. (2021) proponen un modelo basado en la arquitectura de red neuronal híbrida, las cuales son la red neuronal convolucional y LSTM, las cuales emplearon técnicas de reducción de dimensionalidad antes de ser admitidos al clasificador. Por un lado, evidenciaron que en su modelo obtuvo un puntaje de alrededor de 97% de precisión en un conjunto de datos pequeños, por lo que decidieron realizar la prueba de su modelo en un conjunto de datos más grande como PolitiFact, obteniendo como resultado de precisión un 80,62% para el modelo LSTM, 73,29% para el modelo de red neuronal convolucional, 86,14% para el modelo híbrido CNN Y LSTM, 83,81% en el modelo LSTM bidireccional, 88,78% para el modelo de conjunto CNN bidireccional y LSTM con mecanismo de atención, 86,57% en el modelo red neuronal convolucional y LSTM con mecanismo de atención y 86.89% para el modelo ensamblado bidireccional LSTM. Asimismo, para los algoritmos de aprendizaje automático, obtuvieron los resultados de precisión de un 57,58% para la regresión logística y un 58,68% en máquinas de vectores de soporte. Por otro

lado, evidenciaron que el mecanismo de atención utilizado en la red neuronal convolucional y LSTM bidireccional resultó obtener la precisión más alta con un porcentaje del 88,79%.

En conclusión, se han visto diversas investigaciones sobre distintos modelos de inteligencia artificial para detener la propagación de noticias falsas en redes sociales. Por un lado, se han evidenciados diversos modelos basados en inteligencia artificial destacando entre ellos, el modelo de regresión logística por su alto porcentaje de precisión en la detección de noticias falsas en redes sociales. Asimismo, en los últimos años durante la pandemia contra el COVID-19, se ha evidenciado que la propagación de noticias falsas esta afectando cada vez más a la población causando polémicas, enfrentamientos, entre otros. Por lo tanto, el mundo necesita un modelo de detección de noticias falsas con un alto porcentaje de efectividad para la predicción de noticias falsas y reducir su impacto social.

5. **MARCO TEÓRICO**

5.1. Noticias falsas

El concepto o definición de noticias falsas puede tener diferentes significados en diferentes contextos, la cual puede generar confusión en algunos lectores, es por ello por lo que para la presente investigación es necesario aclarar dicha definición. Por un lado, la palabra mencionada “falso” o en inglés “fake” se entiende como algo que no es cierto o aparenta ser verdadero, pero no lo es. Por otro lado, la palabra mencionada “noticias” o en inglés “news”, se refiere a un informe de sucesos que suceden en el presente o que sucederán en el futuro. Por lo tanto, el enunciado “noticias falsas”, indica una información falsa divulgada en forma de sucesos verídicos por medio digital o público. Asimismo, cabe destacar que, en Malasia, está prohibido la divulgación de noticias falsas, la cual si se infringe la legislación obtendrá una sanción de hasta seis años de cárcel(Mokhtar et al. 2019).

5.2. Aprendizaje automático

Es una disciplina científica de la rama de inteligencia artificial, las cuales se crean sistemas que aprenden automáticamente tales como relación de noticias, transcripción de voz a texto, publicidad a los usuarios en cierta red, selección de resultados de búsquedas relevantes e identificación de objetos a través de imágenes(Lecun, Bengio, and Hinton 2015).

5.2.1. Regresión logística

La regresión logística pertenece al tipo de análisis basado en la regresión, la cual es usado para poder predecir el resultado de una variable respecto a las otras variables que son predictoras o independientes. Asimismo, la regresión logística incorpora o incluye la comprobación de hipótesis, en conjunto con las demás evaluaciones realizadas, medidas de aptitud o cálculos. Además, cabe resaltar que la regresión logística podría ser uno de los métodos de explotación en lo que respecta a los datos, los más habituales o usados, ya que cuenta con semejanza o analogía matemática con la sencillez en la que se selecciona de manera automática las variables y la regresión múltiple ordinaria(Desouky Fattoh and Mousa 2021).

5.2.2. Bosque aleatorio

El clasificador de bosque aleatorio está compuesto por una gran variedad de árboles de decisiones, las cuales se desempeñan de manera óptima como un conjunto. Asimismo, por cada árbol individual que se fabrica o construye, realiza uso de la aleatoriedad, para que pueda establecer un bosque de árboles relacionados de forma negativa, la cual su predicción para cualquier árbol será la más confiable. Además, se asigna a cada árbol en el bosque aleatorio una prevención de clase, para lo cual el modelo que se convertirá en predictor será la clase más votada. Cabe destacar, que funciona bien el modelo de bosque aleatorio, ya que los modelos constituyentes individuales se superan gracias a la gran cantidad de árboles correlacionados. Por lo tanto, los bosques aleatorios es un procedimiento o método de vital importancia porque se encuentra basado en el algoritmo de árboles de clasificación y regresión(Desouky Fattoh and Mousa 2021).

5.3. Aprendizaje profundo

Es un proceso de aprendizaje automático basándose en redes neuronales, las cuales resultan muy útiles para la revelación de una o varias estructuras relacionadas con datos de dimensión alta, las cuales son aplicaciones a diversos dominios tales como los negocios en internet, la ciencia, entre otros(Lecun et al. 2015).

5.3.1. Redes neuronales convolucionales

Es un algoritmo de aprendizaje profundo, la cual tiene la capacidad como entrada una imagen ya sea en 2D o 3D, para posteriormente procesarla y al final poder diferenciarlas entre otras imágenes(Lecun et al. 2015).

5.3.2. Redes neuronales concurrentes

Pertenece al grupo de las redes neuronales, las cuales procesan en su entrada una secuencia de elementos a la vez, conteniendo información de elementos pasados basados en historias pasadas, a menudo se usa las redes neuronales recurrentes para tareas relacionadas con la entrada de datos secuenciales, tales como el lenguaje verbal y texto(Lecun et al. 2015).

5.3.2.1. Memoria a corto plazo (LSTM)

La arquitectura de red neuronal recurrente artificial LSTM o memoria a corto plazo, tiene la capacidad de que a través de una larga duración pueda registrar patrones de forma selectiva. Para modelar datos secuenciales puede ser una excelente alternativa, así como también para el estudio de dinámica complejas en lo que respeta al comportamiento humano. Por lo tanto, LSTM o memoria a corto plazo se utiliza mayormente en aplicaciones de múltiples modelos, tales como subtitulado de fotografías tomadas, asimismo también realiza mejoras en las redes neuronales recurrentes o RNN, en contextos híbridos(Desouky Fattoh and Mousa 2021).

5.4. Procesamiento del lenguaje natural (NLP)

El procesamiento de lenguaje natural pertenece a una parte de la inteligencia artificial, la cual es el análisis y procesamiento de un conjunto de datos respecto al lenguaje natural, como puede ser el lenguaje verbal y escrito. Para el caso de la escritura o texto, el propósito es que el texto no estructurado sea cambiado hacia una manera un poco más estructurada, para que pues e pueda obtener información útil cuando se realiza el análisis. Asimismo, el procesamiento del lenguaje natural abarca una serie de metodologías o técnicas para organizar y realizar el procesamiento de los datos

de sentimientos, las cuales pueden ser módulos lingüísticos, vectorización de texto o tokenización(Hybrid, Hawks, and Feature 2021).

5.5. Módulos lingüísticos y de tokenización

Es el proceso por el cual se fraccionan en partes más reducidas el texto, tales como en palabras o expresiones(Hybrid et al. 2021).

5.6. Vectorización de texto

Es un proceso, la cual el texto no estructurado se convierte o transforma en un resultado más organizado, tales que se puedan aplicar los algoritmos de aprendizaje automático con metas a una extracción de conocimientos o de minería, es decir se extraen u obtienen las particularidades numéricas empleando cálculos estadísticos que están basadas en texto o palabras(Hybrid et al. 2021).

5.7. Máquina de vectores de soporte (SVM)

La máquina de vectores de soporte SVM o también conocido por sus siglas SVM, son una agrupación de algoritmos de aprendizaje supervisado, es decir SVM se puede utilizar como un clasificador de pares o también como para la clasificación múltiple. Asimismo, actúa aplicando una técnica para muchas clases en contra de una por cada par de clases. Por lo tanto, el propósito de la máquina de vectores de soporte es identificar o reconocer las instancias, la cual tiene dependencia a una ecuación lineal, así también se puede ejecutar con una función de kernel, una clasificación no lineal(Desouky Fattoh and Mousa 2021).



6. HIPÓTESIS

7. METODOLOGÍA

8. CRONOGRAMA DE TRABAJO

9. PRESUPUESTO

10. BIBLIOGRAFÍA

Ali, Hassan, Muhammad Suleman Khan, Amer AlGhadhban, Meshari Alazmi, Ahmad Alzamil, Khaled Al-utaibi, and Junaid Qadir. 2021. "Analyzing the Robustness of Fake-News

- Detectors under Black-Box Adversarial Attacks.” *IEEE Access*. doi: 10.1109/ACCESS.2021.3085875.
- Desouky Fattoh, Ibrahim E. L., and Farid Ali Mousa. 2021. “Fake News Detection Based on Word and Document Embedding Using Machine Learning Classifiers.” *Journal of Theoretical and Applied Information Technology* 99(8):1891–1901.
- Hansrajh, Arvin, Timothy T. Adeliyi, and Jeanette Wing. 2021. “Detection of Online Fake News Using Blending Ensemble Learning.” *Scientific Programming* 2021. doi: 10.1155/2021/3434458.
- Hybrid, Utilizing, Harris Hawks, and Based Feature. 2021. “SS Symmetry Intelligent Detection of False Information in Arabic Tweets Machine Learning Models.”
- Jiang, Tao, Jian Ping Li, Amin Ul Haq, Abdus Saboor, and Amjad Ali. 2021. “A Novel Stacking Approach for Accurate Detection of Fake News.” *IEEE Access* 9:22626–39. doi: 10.1109/ACCESS.2021.3056079.
- Lecun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. “Deep Learning.” doi: 10.1038/nature14539.
- Mahabub, Atik. 2020. “A Robust Technique of Fake News Detection Using Ensemble Voting Classifier and Comparison with Other Classifiers.” *SN Applied Sciences* 2(4):1–9. doi: 10.1007/s42452-020-2326-y.
- Mahlous, Ahmed Redha, and Ali Al-Laith. 2021. “Fake News Detection in Arabic Tweets during the COVID-19 Pandemic.” *International Journal of Advanced Computer Science and Applications* 12(6):778–88. doi: 10.14569/IJACSA.2021.0120691.
- Mokhtar, Muhammad Syahmi, Yusmadi Yah Jusoh, Novia Admodisastro, Norainiche Pa, and Amru Yusrin Amruddin. 2019. “Fakebuster: Fake News Detection System Using Logistic Regression Technique in Machine Learning.” *International Journal of Engineering and Advanced Technology* 9(1):2407–10. doi: 10.35940/ijeat.A2633.109119.
- Zervopoulos, Alexandros, Aikaterini Georgia, Alvanou Konstantinos, and Bezas Asterios. 2021. “Deep Learning for Fake News Detection on Twitter Regarding the 2019 Hong Kong

Protests.” *Neural Computing and Applications* 0. doi: 10.1007/s00521-021-06230-0.

11. **ANEXOS**