



**Universidad de las Fuerzas Armadas “ESPE”**  
**Lectura y Escritura de Textos Académicos**

**Resumen Estado del Arte**

**TEMA:**

“Identificación de hechos y opiniones en redes sociales digitales utilizando técnicas de aprendizaje automático “

**Integrantes:**

Freddy Rodríguez

**Tutora:** Ing Eleana Inés Jerez Villota, MSc

## Palabras clave

Misinformation, disinformation, and true information.

dissemination in social networks.

## Cadena de busqueda

Dissemination of information in social networks **AND** misinformation in social networks **AND** machine learning

### A. Preguntas a investigar

- ¿Qué tipo de información se propaga intencionalmente en redes sociales?
- ¿Cuáles son las herramientas que permiten extraer información desde las redes sociales digitales?

### B. Búsqueda primaria

Como primer punto se estableció una búsqueda para elaborar la cadena

	TEMA GC
A-GC1	Automatic Rumour Detection Model on Social Media
A-GC2	A Multiple Feature Category Data Mining and Machine Learning Approach to Characterize and Detect Health Misinformation on Social Media
A-GC3	Misinformation Detection on social media: Challenges and the Road Ahead
A-GC4	Dataset for Arabic Fake News
A-GC5	Using Deep Learning to Identify Linguistic Features that Facilitate or Inhibit the Propagation of Anti- and Pro-Vaccine Content on Social Media
A-GC6	User Behavior Analysis based on their Social Media Interaction
A-GC7	An Adaptive Approach for Fake News Detection in Social Media: Single vs Cross Domain
A-GC8	Combating Misinformation Dissemination through Verification and Content Driven Recommendation
A-GC9	A Multitask Learning Approach for Fake News Detection: Novelty, Emotion, and Sentiment Lend a Helping Hand

### C. Resumen de los artículos seleccionados

#### - Automatic Rumour Detection Model on social media

El presente artículo se enfoca en la red social digital Twitter, que se ha convertido en un lugar popular para realizar opiniones que generan caos dentro de la sociedad por la información que se comparte. Este estudio se caracteriza

por ser descriptivo y recolectar tweets para entregarlos en marco TF-IDF para generar un conjunto de características compuestas. El mecanismo potencial ayudará a detectar información errónea lo más rápido posible para contrarrestar la difusión de rumores y generar una profunda confianza de los usuarios en los sitios de redes sociales.

- **A Multiple Feature Category Data Mining and Machine Learning Approach to Characterize and Detect Health Misinformation on social media**

En este artículo, se describe la de infiltración errónea sobre la salud como un proceso dinámico de difusión en las redes sociales, así como propiedades basadas en el contenido. Utilizando las discusiones sobre el Zika en Twitter en 2016 como marco de investigación, identificamos los 264 tuits de desinformación más influyentes y relacionamos 455 tuits con información real. Desarrollamos un algoritmo para inferir la red de propagación de información iterando cada tweet y obteniendo nueve métricas de red. Luego, aproximamos la propagación de la información como una señal de proceso de Poisson no homogénea (NHPP). Luego obtuvimos 40 características distintivas para caracterizar cada NES. Para las funciones basadas en contenido, utilizamos consultas lingüísticas y conteo de palabras y vectorización de documentos para obtener 63 y 50 funciones para cada tweet, respectivamente.

- **Misinformation Detection on social media: Challenges and the Road Ahead**

Cada vez es más difícil lidiar con el volumen, la variedad, la velocidad y la precisión de la desinformación (como contenido de noticias falsas, publicaciones falsas e imágenes/videos falsos) que se difunde desde varias plataformas en línea. En este documento, proporcionamos una descripción general de los métodos existentes basados en el aprendizaje automático para detectar la desinformación y el ocultamiento de información, y discutimos las amenazas y limitaciones actuales de estos enfoques. Con base en la discusión, identificamos algunas contramedidas posibles.

- **Using Deep Learning to Identify Linguistic Features that Facilitate or Inhibit the Propagation of Anti- and Pro-Vaccine Content on Social Media**

Gracias a las redes sociales se han bajado las barreras para publicar contenido. Si bien el empoderamiento personal tiene resultados positivos, también permite que los actores malintencionados creen información errónea y la difundan entre usuarios aleatorios de la plataforma. Si bien las señales de reputación y las banderas de verificación de hechos brindan filtros y señales útiles a los lectores, y los algoritmos y los moderadores humanos ayudan a detener la difusión de información errónea, ninguna de estas intervenciones aborda la arquitectura básica de las redes sociales que permite que se propague la información errónea. Las celebridades de las redes sociales, especialmente los usuarios de

YouTube, han incursionado en los canales principales, como las redes de transmisión y de cable, en los últimos años, provocando grandes cambios en la forma en que las personas se comportan en las redes sociales. Algo está afectado.

- **An Adaptive Approach for Fake News Detection in Social Media: Single vs Cross Domain**

El uso generalizado de los medios tradicionales en las plataformas de información en línea amplifica la difusión de noticias falsas. Los métodos basados en aprendizaje automático supervisado son muy utilizados para detectar noticias falsas en las redes sociales. Sin embargo, en escenarios de datos de dominios cruzados, el rendimiento de dichos modelos se degradará. En este estudio, demostramos empíricamente que el rendimiento del modelo es específico del dominio e independiente del caso. Para realizar esta investigación, extrajimos tweets basados en la crisis de Afganistán y desarrollamos un conjunto de datos que llamamos "FakeBan". El país ha sido testigo de una repentina proliferación de desinformación, mal utilizada como munición por varios actores, con consecuencias inquietantes y de largo alcance. Recientemente elegimos estudiar Afganistán y experimentamos con tres áreas diferentes de noticias falsas: crisis nacional, atención médica y política. Ya existen varios conjuntos de datos de alto nivel en los dominios de salud y políticas. Sin embargo, tomó mucho tiempo crear un conjunto de datos etiquetados basado en la reciente crisis nacional. Presentamos una técnica adaptativa de detección de noticias falsas capaz de seleccionar patrones por dominio (único o multidominio), resolviendo así el desafiante problema de la gran cantidad y diversidad de información disponible en las redes sociales.

## **Resumen**

Las redes sociales digitales como: ResearchGate, una red social de investigadores cuyo enfoque principal es desarrollar ciencia de cualquier disciplina; Twitter, utilizada para publicar información en tiempo real acerca de temáticas relacionadas con problemas políticos, económicos, sociales, etc.; Facebook, aquí las personas propagan información mediante texto, fotos o videos en tiempo real acerca de una gran variedad de temáticas; Instagram, una red social utilizada para compartir fotografías y videos; TikTok, el enfoque principal de esta red social es grabar, editar y compartir vídeos que se publican en tiempo real con varias temáticas relacionadas con tecnología, política, económica y polémica social ; LinkedIn, considerada como una red social digital para profesionales, que muestra un pequeño curriculum de las personas; se han convertido en objeto de análisis de muchos investigadores. Por ejemplo, algunos estudios se enfocan en definir modelos de propagación de información, otros en mitigar la información que se propaga identificando a los nodos semilla y algunos en el análisis de la desinformación. Además, en la literatura académica se ha encontrado una clasificación de la información con los términos en inglés:

“misinformation”, “disinformation” y “true information”. Sin embargo, no se han encontrado trabajos en los que se identifique tipos de información basados en la argumentación que se propagan intencionalmente en las redes sociales digitales como los hechos y las opiniones utilizando técnicas de aprendizaje automático. La identificación de estos tipos de información, más específica y menos generalizada que la que se ha encontrado, nos permitirá analizar de mejor manera el comportamiento de la información que se propaga en las redes sociales digitales, considerando que en muchas ocasiones puede causar caos y pánico colectivo en la población.

## REFERENCIAS

Ahvanooey, M. T., Zhu, M. X., Mazurczyk, W., Choo, K. K. R., Conti, M. & Zhang, J.

(2022). Misinformation Detection on Social Media: Challenges and the Road

Ahead. *IT Professional*, 24(1), 34-40.

<https://doi.org/10.1109/mitp.2021.3120876>

Bharti, M. & Jindal, H. (2020). Automatic Rumour Detection Model on Social Media.

*2020 Sixth International Conference on Parallel, Distributed and Grid*

*Computing (PDGC)*. <https://doi.org/10.1109/pdgc50313.2020.9315738>

Pushkar, P. & Mittal, U. (2022). User Behavior Analysis based on their Social Media

Interaction. *2022 3rd International Conference on Intelligent Engineering and*

*Management (ICIEM)*. <https://doi.org/10.1109/iciem54221.2022.9853113>

Rastogi, S., Gill, S. S. & Bansal, D. (2021). An Adaptive Approach for Fake News

Detection in Social Media: Single vs Cross Domain. *2021 International*

*Conference on Computational Science and Computational Intelligence (CSCI)*.

<https://doi.org/10.1109/csci54926.2021.00280>

Safarnejad, L., Xu, Q., Ge, Y. & Chen, S. (2021). A Multiple Feature Category Data

Mining and Machine Learning Approach to Characterize and Detect Health

Misinformation on Social Media. *IEEE Internet Computing*, 25(5), 43-51.

<https://doi.org/10.1109/mic.2021.3063257>

