MONSTER: Una arquitectura para la detección de noticias falsas

Marta Fuentes-García
Fundación I+D del Software Libre (Fidesol)
Avda. de la Innovación, 1
Ed. BIC-CEEI 18016 - Granada
mfuentes@fidesol.org

***Resumen*—Las noticias falsas han existido durante años. Hay registros que incluso se remontan a la Antigua Roma. Desde entonces, nuestra sociedad ha sufrido los efectos negativos derivados de las noticias falsas, que se expanden a un ritmo vertiginoso. Esta expansión ha cobrado especial importancia en los últimos años, con el auge de las redes sociales. Por esta razón, es de especial interés recuperar la confianza de los usuarios en la información que reciben. Así, la detección y prevención de la difusión de noticias falsas es realmente importante para evitar efectos negativos como la polarización y el sesgo en la forma de pensar de la humanidad. En los últimos años se ha hecho un gran esfuerzo por parte de la ciencia y la industria para proponer soluciones que ayuden a reducir el número de noticias falsas y su efecto en la sociedad. Estos esfuerzos se han centrado tradicionalmente en la aplicación de técnicas de aprendizaje automático para detectar contenido falso. Por otra parte, con la expansión del *blockchain* existe una nueva tendencia en el uso de esta tecnología para asegurar tanto la integridad como la trazabilidad del contenido digital. En este trabajo, presentamos *Monster*, un enfoque híbrido que combina *blockchain* y aprendizaje automático en una única arquitectura para la detección de noticias falsas.**

***Index Terms*—Noticias falsas, Blockchain, Aprendizaje automático, Detección de noticias falsas, Identidad digital**

I. Introducción

Estamos ante una nueva revolución de la información [1] que permite acceder contenido en cualquier momento y lugar. Al mismo tiempo, este contenido puede ser publicado por cualquiera. Sin embargo, la fuente de esta información, así como el contenido en sí mismo, no siempre están verificados y/o contrastados. El término noticia falsa (en inglés, *Fake News*) se refiere a aquella noticia que está contrastada como falsa y que confunde a los lectores [2]. Dicha noticia puede haber sido redactada intencionadamente o ser fruto de un rumor o bulo. En algunas ocasiones se pueden considerar como noticias falsas aquellas que son satíricas, es decir, cuyo contenido se sabe que es falso y tiene como único objetivo el entretenimiento del público que las consume [3–5].

Las noticias falsas influyen en la opinión de las personas, por lo que pueden tener serias consecuencias en la sociedad, como la polarización en los ideales o la desconfianza por parte de la población [2, 6–9]. Esta problemática es especialmente relevante en áreas relacionadas con la sanidad [3, 4] o la gestión de emergencias [10–12]. En este sentido, redes sociales como Twitter se postulan a la vez como medio rápido de información y comunicación ante noticias de última hora, pero también como principal foco de desinformación y expansión de noticias falsas y rumores [2, 4, 5, 7–11]. Las redes sociales hacen que sea extremadamente fácil compartir contenido que puede ser erróneo, con el consecuente riesgo para la salud cuando dicho contenido se refiere a enfermedades y temas sanitarios. En este sentido, la situación actual derivada del COVID-19 ha disparado una ola de noticias falsas, rumores y bulos que pone en riesgo la seguridad de la sociedad [13, 14]. Este problema empeora cuando celebridades o incluso políticos son los que difunden información falsa [15, 16]. Uno de los ejemplos más conocidos es la recomendación del uso de Hidroxicloroquina como tratamiento para el COVID-19, que ha sido defendida en varias ocasiones por el ya ex-presidente de los Estados Unidos, Donald Trump, pero que no está aceptada por la Organización Mundial de la Salud [17, 18].

En un estudio realizado por Waszak *et al* en 2018 [19] se observó que casi la mitad de los enlaces compartidos con mayor frecuencia contenía noticias catalogadas como *Fake News*. Otro problema relacionado con las noticias falsas es la conocida como ”Post-Verdad”, consistente en la distorsión de la realidad relacionada con la falta de objetividad en los hechos y en la búsqueda de noticias o fuentes afines al pensamiento particular de cada persona [2–4, 7, 20].

Un estudio publicado por Oh *et al* [10] demostró que existen dos variables clave que explican el comportamiento en situaciones extremas. Dichas variables son la ”ansiedad” y la ”ambigüedad de la información”. Según este mismo estudio, una forma de reducir el impacto de estas variables, especialmente de la ansiedad, es obtener la información de fuentes creíbles. Una fuente confiable o creíble es aquella que se dedica a la publicación de noticias veraces (como por ejemplo periódicos de tirada nacional) o un organismo público (como por ejemplo la Organización Mundial de la Salud o la Policía Nacional) [21]. Dichas fuentes publican información original, real y cuyo contenido puede ser verificado como no alterado. Además, existen ciertas páginas web cuyo principal objetivo es desmentir rumores y noticias falsas [22–25]. Por otro lado, en España hay ciertos criterios que pueden permitir distinguir una noticia real de una falsa [26], aunque éstos no son aplicables cuando se trata de noticias difundidas a través de redes sociales y podrían no ser extensibles a noticias procedentes de otros países.

En este trabajo presentamos ***Monster*** (del inglés, *MOnitoring NewS for Trust Enhanced Recovering*), un enfoque que combina aprendizaje automático y *blockchain* para la detección de noticias falsas. Se trata de un trabajo teórico en el que presentamos el concepto de arquitectura modular y evaluamos las ventajas e inconvenientes de su implementación. El objetivo de ***Monster*** es proporcionar las siguientes características: *i)* **detectar y notificar** si el contenido de una

noticia es potencialmente falso, *ii)* disponer de una lista de **fuentes confiables** que puedan validar el contenido publicado, *iii)* asegurar la **inmutabilidad y robustez** del contenido de las noticias, y *iv)* garantizar la **trazabilidad y no repudio** de las noticias. Se pretende así mejorar y recuperar la confianza de los usuarios en el contenido que consumen.

El resto del artículo está organizado como sigue: la Sección II revisa la literatura existente relacionada con la detección de noticias falsas. En la Sección III se presenta una solución novedosa para la detección de noticias falsas. Finalmente, las conclusiones derivadas de este trabajo se presentan en la Sección IV.

II. Trabajo relacionado

Existen numerosas propuestas para la detección de *Fake News*. En 2018, el Departamento de Seguridad de Homeland presentó un libro blanco para la contención de noticias falsas en redes sociales durante desastres y casos de emergencia [12]. En este libro se recogen algunas pautas como que las respuestas de emergencia deberían incluir planes para contener rumores, desinformación y noticias falsas en casos de desastres. Desde Facebook se ha realizado una serie de propuestas para acabar con la difusión de *Fake News* [27]. Entre otras medidas, se indica que proporcionar un mayor contexto e información puede hacer más fácil la tarea de informar y decidir sobre la veracidad o no de una noticia [27]. Por su parte, aplicaciones como Whatsapp restringen la cantidad de mensajes que un usuario puede compartir con el mismo contenido [28], además proporciona una serie de consejos para que los usuarios aprendan a distinguir noticias falsas [29].

Otra gran tendencia es la de aplicar algoritmos de detección automática. En [30] se propone un algoritmo de verificación de hechos, basado en la técnica periodística. Básicamente consiste en identificar **quién**, **dónde**, **cuándo**, así como la **temática** y el **hecho** dentro de la noticia. Sin embargo, la mayoría de propuestas de la literatura están basadas en la aplicación de aprendizaje automático (ML, del inglés, *Machine Learning*). Algunas de las técnicas que destacan son el procesamiento del lenguaje natural (NLP, del inglés, *Natural Language Processing*) [28], SVM (del inglés *Support Vector Machine*) [11, 28, 31–33], Naïve-Bayes [11, 28, 31, 34], regresión [5, 28, 31], árboles de decisión [28, 31] o redes neuronales [7–9, 11, 19, 28, 31, 35, 36]. Se puede encontrar una revisión completa sobre la aplicación de ML para la detección de noticias falsas en [37]. Por lo general, todas las soluciones se relacionan con el análisis de características lingüísticas del texto. Otros autores apuestan por la combinación de distintos algoritmos para optimizar la capacidad de detección [38, 39]. En el caso de detección de noticias falsas en redes sociales, la mayoría de los modelos se basan además en el estudio del comportamiento asociado a usuarios (perfiles que suelen interesarse en noticias falsas y perfiles que suelen interesarse en noticias veraces), así como en el impacto que causa una noticia medido en cantidad de 'me gusta'.

Por otra parte, la tecnología *blockchain* surgió hace algo más de una década para dar soporte a las criptomonedas *bitcoins* [40]. Desde entonces han sido múltiples las aplicaciones de *blockchain* que van más allá de las finanzas, debido a que se caracteriza por ser resistente a falsificaciones y permitir la verificación multinodo de contenido [2, 6, 41, 42]. Así, ningún nodo puede modificar de forma aislada el contenido previo de los bloques que ya han sido modificados. Para modificarlos deben estar de acuerdo al menos el 50 % de los nodos [41]. Para ello, *blockchain* utiliza una estructura de almacenamiento distribuido, criptografía, un algoritmo de consenso y contratos inteligentes, entre otros [6]. Es por ello que algunos trabajos proponen la aplicación de *blockchain* para la verificación de contenido y trazabilidad de noticias [2, 6, 9, 20, 43]. De esta forma, se propone la identificación de la fuente que publica una noticia gracias a las identidades digitales distribuidas, así como el uso de la propiedad de inmutabilidad de *blockchain* para detectar modificaciones en el contenido de las noticias [6, 20].

III. ***Monster***: ML y *blockchain* para la detección de *Fake News*

En este trabajo, proponemos ***Monster*** (del inglés, *MOnitoring NewS for Trust Enhanced Recovering*), un sistema basado en la monitorización que combina *blockchain* y ML para la detección de noticias falsas.

Por una parte, *blockchain* se utiliza con un propósito doble: *i)* asegurar la inmutabilidad y robustez del contenido y *ii)* asegurar la trazabilidad y no repudio de la noticia. Recordemos que *blockchain* garantiza que el contenido no se puede modificar sin el acuerdo de un número mínimo de nodos y, si finalmente se modifica, esta alteración del contenido original es detectable [2, 6, 9, 41, 42, 44]. Además, las identidades digitales descentralizadas permiten identificar las fuentes confiables y rastrear el origen (y/o publicador) de la noticia, así como de cada uno de sus componentes (texto, imagen, vídeo y audio) [42, 44–46].

Por otra parte, los algoritmos de ML permiten tratar con grandes cantidades de datos e identificar así patrones que no serían detectables por el ser humano de manera directa. Esto hace que estas técnicas sean ampliamente exploradas en los trabajos relacionados con la detección de noticias falsas. Así, la mayoría de los algoritmos de aprendizaje automático utilizados para la detección de noticias falsas son de *clasificación*[1] y *clustering*[2]. La tendencia más extendida en la literatura es la de aplicar clasificación, de forma que los autores entrenan un modelo con noticias etiquetadas como '**verdadera**' o '**falsa**', aunque a veces hacen distinciones en el grado de falsedad (p. e. '**parcialmente cierta**' o '**mayoritariamente cierta**') [5, 7, 8, 11, 31, 33, 34, 37]. Además, existen herramientas que permiten detectar noticias falsas de forma automática [47–49]. Los autores de [47] proponen un complemento que se instala en el navegador y permite eliminar de los resultados de búsqueda sitios que proporcionan noticias falsas. Por otra parte, Pantech ha implementado una solución en Python que aplica TF-IDF (NLP) para detectar noticias falsas [49]. De forma similar, en [48] se propone otra alternativa para la detección de noticias falsas, también basada en la técnica de TF-IDF e implementada en Python.

[1] La **clasificación** es una técnica de aprendizaje supervisado y, por tanto, necesita que los datos de entrenamiento estén etiquetados con las clases a identificar. Los algoritmos de clasificación asignan la clase que mejor encaja con los datos de entrada.

[2] El *clustering* es una técnica de aprendizaje no supervisado y, por tanto, no necesita que los datos de entrenamiento estén etiquetados. Los algoritmos de *clustering* identifican grupos (*clusters*) con características similares en los datos de entrada.

En este artículo proponemos un enfoque diferente, basado en la *detección de anomalías*. La **detección de anomalías** tiene como objetivo detectar cuando los datos no siguen el mismo comportamiento que los utilizados para la construcción del modelo (entrenamiento) [50]. En este caso, el objetivo es dividir los datos correspondientes a las noticias en dos grupos: verdaderas y no verdaderas. Al igual que para el resto de técnicas de ML, es importante que el conjunto de datos de entrenamiento esté balanceado. Sin embargo, en la detección de anomalías éste únicamente estará formado por noticias que han sido contrastadas como ciertas. De esta forma, las noticias que no se ajustan al modelo generado se consideran anómalas y, por tanto, se marcan como potencialmente falsas.

Para generar el conjunto de datos de entrenamiento proponemos capturar noticias de un conjunto de medios nacionales (periódicos y páginas de verificación). Es importante elegir medios de ideología variada para evitar posibles sesgos. Además, en este conjunto de datos se incluirán publicaciones de redes sociales llevadas a cabo por estos mismos medios, por fuentes oficiales (como Policía Nacional), y por periodistas y programas informativos.

En cuanto a la estrategia de aprendizaje, optaremos por clasificadores de una sola clase, ya que en este tipo de algoritmos sólo hay dos posibilidades: los datos evaluados (en nuestro caso, las noticias) pertenecen a la clase (noticias verdaderas) o no. Así, si se utiliza aprendizaje supervisado, la etiqueta del conjunto de entrenamiento será 'verdadera'. En cuanto a la salida, las noticias se etiquetarán (independientemente de si el algoritmo es supervisado o no) como: '**potencialmente verdaderas**' (si encajan con los datos de entrenamiento) y '**potencialmente falsas**' (si no encajan con los datos de entrenamiento). Para elegir la estrategia final de detección se llevará a cabo un estudio comparativo en el que se evaluará el rendimiento de distintos algoritmos de detección de anomalías dentro de la arquitectura propuesta.

Actualmente ***Monster*** se encuentra en una fase preliminar y está siendo desarrollado de forma colaborativa como parte de los proyectos ***EGIDA***[3], ***IASEC*** y ***ASAC blockchain***. La Tabla I resume las principales características de ***Monster***, considerando tanto las ventajas ('✓') como posibles inconvenientes ('✗') de esta propuesta. Nótese que estos posibles inconvenientes son problemas que podrían surgir durante la implementación y/o uso efectivo de esta solución.

La Fig. 1 muestra el concepto de arquitectura propuesta para ***Monster***. Esta arquitectura se compone principalmente de cuatro módulos: **RIS**, **FEX**, **MON**, y **TRUST**; que se describen en los próximos párrafos.

Módulo RIS (del inglés Reading, Integrating and Saving): Este módulo lleva a cabo la captura de datos de una o más fuentes. Se utiliza tanto para la construcción del modelo como para la monitorización de nuevas noticias. Se alimenta de bases de datos disponibles, utilizando API para recuperar las noticias. Además, tiene acceso a medios de difusión contrastados (como periódicos o cuentas verificadas de redes sociales) así como a redes sociales. Adicionalmente, RIS se encarga de unificar y almacenar los datos capturados en la base de datos homónima.

[3]https://egidacybersecurity.com/

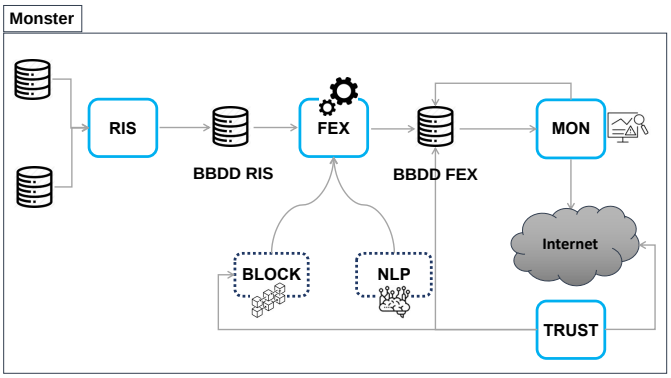

Figura 1. Concepto de arquitectura para el sistema ***Monster***.

Módulo FEX (del inglés Feature Extraction): Este módulo lee los datos en crudo de la base de datos RIS y realiza una **extracción de características** a partir de esos datos. Para ello, depende de módulos y subsistemas adicionales. Estas características, en combinación con algunas de las que se obtienen directamente mediante el módulo RIS, se almacenan en una nueva base de datos, FEX. Algunas de las características que se pueden obtener directamente de los datos son las siguientes:

- **Multimedia**: {0, 1}. Tanto en periódicos como en redes sociales, puede existir contenido multimedia (1) o no (0).
- **Número de reacciones**. Las noticias falsas suelen generar un elevado número de reacciones. Estas reacciones pueden ser, por ejemplo, comentarios o número de 'me gusta'. El objetivo de esta característica es ayudar a obtener estadísticas que podrían ayudar a la detección de una noticia falsa.

Los módulos adicionales forman parte de dos subsistemas que se encargan de:

- **NLP**. El objetivo de este subsistema es obtener características relacionadas con el texto. En concreto, el módulo de NLP permite extraer características de:
  - **Conteo**. Es necesario llevar a cabo un procesamiento del texto para extraer las características relacionadas con estadísticas del texto. Por ejemplo: extensión de la noticia (número de caracteres y/o número de palabras).
  - **Temática**. Permite clasificar la noticia en cuanto al área de interés. Es posible que unos temas sean más tendentes a la propagación de noticias falsas. El objetivo de esta característica es ayudar a obtener estadísticas que podrían ayudar a la detección de una noticia falsa. Para ello, es necesario llevar a cabo un procesamiento del texto que permita detectar la temática de la noticia. Además, es necesario definir un listado de categorías y emplear técnicas de clasificación. Un ejemplo de categorías podría ser: *política*, *deportes*, *salud*, *famosos*, y *otros*.
  - **Sentimientos**. Permite clasificar la noticia en cuanto al sentimiento global transmitido. Por lo general, una noticia real, publicada por un medio confiable, debería ser clasificada con un sentimiento 'neutro'. Sin embargo, una noticia falsa tiende a la polarización de sentimientos, es decir, se puede clasificar como

Tabla I
CARACTERÍSTICAS DE ***Monster*** PARA LA DETECCIÓN DE NOTICIAS FALSAS. EL SÍMBOLO ✓ REPRESENTA LAS VENTAJAS MIENTRAS QUE ✗ REPRESENTA LOS INCONVENIENTES DE LA PROPUESTA.

✓	✗
Fuente de publicación confiable identificable	Se necesita verificación y registro previo de fuentes confiables en la red de *blockchain*
Trazabilidad	
Integridad y robustez de contenido	
Detección de alteración de contenido	
Válido para contenido multimedia	El contenido es complejo (texto + multimedia)
Válido para redes sociales	Se necesita un conjunto balanceado de noticias previamente contrastadas como ciertas
Se consideran las características del lenguaje	Se necesitan conjuntos de datos independientes para tener en cuenta distintos idiomas
Se considera el tema de las noticias	
Se considera el análisis de sentimientos	
Se permite la verificación de hechos	Verificación de hechos no automática (requiere fuentes verificadas como fiables)
Actualización periódica del modelo	

'positiva' o 'negativa'. El objetivo de esta característica es ayudar a obtener estadísticas que podrían ayudar a la detección de una noticia falsa. Para ello, es necesario llevar a cabo un procesamiento del texto y, posteriormente, un análisis de sentimientos para catalogar cada noticia como 'positiva', 'negativa', o 'neutra'.

- **Blockchain**. El objetivo de este subsistema es llevar a cabo las acciones relacionadas con la red de *blockchain*. Este módulo pretende proporcionar trazabilidad de la noticia y los publicadores, así como integridad de contenido, por lo que es necesario que las fuentes confiables se encuentren almacenadas en la red de *blockchain*, así como el uso de identidades digitales para identificar si la noticia procede o no de una fuente confiable. También es necesario que las noticias (tanto el texto como los elementos multimedia) se encuentren registrados en la red de *blockchain* para poder detectar si han sido modificados o no. Además, el subsistema de blockchain permite derivar las siguientes características que se utilizan para la construcción del modelo:
  - **Fuente confiable**: $\{0, 1\}$. Si una noticia ha sido publicada y/o compartida por fuentes confiables, es menos probable que no sea falsa. Nótese que, en este caso, el valor 0, no significa que no sea confiable, sino que no está registrada en la cadena de *blockchain* como tal. En cambio, el valor 1 sí es sinónimo de fuente confiable registrada en la cadena de *blockchain*.
  - **Elemento multimedia modificado**: $\{0, 1\}$. Si una noticia contiene un elemento multimedia (vídeo, fotografía o audio), este puede haber sido modificado (1) o no (0). Las noticias falsas suelen contener contenido modificado. Sin embargo, que el contenido multimedia haya sido modificado no implica que la noticia sea necesariamente falsa.
  - **Elemento texto modificado**: $\{0, 1\}$. El texto de una noticia puede haber sido modificado (1) o no (0). Las noticias falsas suelen contener contenido modificado. Sin embargo, que el texto haya sido modificado no implica que la noticia sea necesariamente falsa.
  - **Enlace a fuente confiable**: $\{0, 1\}$. Una noticia puede contener enlace a fuentes confiables (1) o no (0). Si una noticia contiene enlaces a fuentes confiables, es menos probable que sea falsa.

Módulo MON (del inglés, MONitoring): En este módulo se distinguen en dos fases:

1. **Construcción del modelo**. Se lleva a cabo a partir de datos históricos. Dado que se aplica detección de anomalías, el modelo se genera utilizando noticias que han sido contrastadas como ciertas.
2. **Monitorización** de noticias. Se encarga de detectar si las nuevas noticias siguen el mismo patrón que las que se utilizaron para crear el modelo. Si la nueva noticia no sigue el modelo creado a partir de noticias fiables, entonces se considera anormal (anómala), se etiqueta como '***potencialmente falsa***' y se genera una alarma. Si, por el contrario, sigue el modelo generado en la primera fase, entonces se etiqueta como '***potencialmente verdadera***'.

Módulo TRUST: Este módulo permite **verificar** si una noticia es verdadera o no. Esta acción es llevada a cabo por fuentes fiables, previamente registradas en la red de *blockchain*. Cuando una noticia es validada (bien sea verdadera o falsa), los resultados se actualizan en la base de datos FEX y en la red de *blockchain*. El modelo se actualiza de forma periódica para incluir el *feedback* recibido del módulo TRUST, de forma que el número de falsos positivos y negativos se reduce, mejorando así el rendimiento y la fiabilidad del sistema completo.

IV. CONCLUSIONES

En este artículo proponemos ***Monster***, una solución novedosa para la detección de noticias falsas. ***Monster*** se beneficia del uso del aprendizaje automático y *blockchain*, combinándolos en una única arquitectura para satisfacer los siguientes objetivos: *i)* detectar y notificar si el contenido de una noticia es potencialmente falso o no, *ii)* proporcionar una lista de fuentes confiables que puedan validar el contenido de una noticia, *iii)* asegurar la inmutabilidad y la robustez del contenido de una noticia, y *iv)* garantizar la trazabilidad y el no repudio del contenido de una noticia.

Las características proporcionadas por ***Monster*** se resumen en una tabla que incluye las ventajas y los inconvenientes que podrían aparecer durante la implementación y despliegue del sistema. Esta propuesta se encuentra todavía en fase de desarrollo, como parte de un trabajo colaborativo entre los proyectos ***EGIDA***, ***IASEC*** y ***ASAC blockchain***. En los próximos meses llevaremos a cabo experimentos en un entorno controlado de laboratorio con un conjunto de datos de

prueba. Nuestro objetivo final es desplegar ***Monster*** como una solución que pueda ser utilizada por organizaciones y compañías para mitigar el efecto negativo de la difusión de noticias falsas.

Agradecimientos

Este trabajo está financiado por las Ayudas Cervera para Centros Tecnológicos del Centro español para el Desarollo de Tecnología Industrial (CDTI) bajo el proyecto EGIDA (CER-20191012).

Referencias

[1] J. S. Nye, "The Information Revolution and Soft Power," *Current History*, vol. 113, no. 759, 2014.

[2] A. Qayyum, J. Qadir, M. U. Janjua, and F. Sher, "Using Blockchain to Rein in the New Post-Truth World and Check the Spread of Fake News," *IT Professional*, vol. 21, no. 4, pp. 16–24, 2019.

[3] P. M. Waszak, W. Kasprzycka-Waszak, and A. Kubanek, "The spread of medical fake news in social media -The pilot quantitative study," *Health Policy and Technology*, vol. 7, no. 2, pp. 115 – 118, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2211883718300881

[4] M. Viviani and G. Pasi, "Credibility in social media: opinions, news, and health information - a survey," *WIREs Data Mining and Knowledge Discovery*, vol. 7, no. 5, p. e1209, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1209

[5] E. Tacchini, G. Ballarin, M. L. D. Vedova, S. Moret, and L. de Alfaro, "Some Like it Hoax: Automated Fake News Detection in Social Networks," in *Proceedings of the Second Workshop on Data Science for Social Good (SoGood), Skopje, Macedonia, 2017 (arXiv:1704.07506v1)*, vol. 1960, 2017.

[6] W. Shang and M. Liu and W. Lin and M. Jia, "Tracing the Source of News Based on Blockchain," in *IEEE ACIS 17th International Conference on Computer and Information Science (ICIS)*, 2018.

[7] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," *SIGKDD Explor. Newsl.*, vol. 19, no. 1, pp. 22–36, 2017. [Online]. Available: https://doi.org/10.1145/3137597.3137600

[8] N. Ruchansky, S. Seo, and Y. Liu, "CSI: A Hybrid Deep Model for Fake News Detection," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, ser. CIKM '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 797–806. [Online]. Available: https://doi.org/10.1145/3132847.3132877

[9] T. W. Jing and R. K. Murugesan, "A theoretical framework to build trust and prevent fake news in social media using blockchain," in *Recent Trends in Data Science and Soft Computing*, F. Saeed, N. Gazem, F. Mohammed, and A. Busalim, Eds. Cham: Springer International Publishing, 2019, pp. 955–962.

[10] O. Oh, K. Kwon, and H. Rao, "An exploration of social media in extreme events: Rumor theory and twitter during the HAITI earthquake 2010," in *International Conference on Information Systems, ICIS 2010*, 2010, cited By 86. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84870963741&partnerID=40&md5=a6f0a1cbe56d3e2901be541efbdaa460

[11] S. Aphiwongsophon and P. Chongstitvatana, "Detecting fake news with machine learning method," in *2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, July 2018, pp. 528–531.

[12] Homeland-Security, "Countering false information on social media in disasters and emergencies," Homeland Security, https://cutt.ly/3tj2W9m, Tech. Rep., 2018, [Online, accessed on 17/03/2020].

[13] EUROPOL, "COVID-19: Fake News," https://www.europol.europa.eu/covid-19/covid-19-fake-news, 2020, [Online, accessed on 07/10/2020].

[14] PWC, "How fake news has exploited COVID-19," https://cutt.ly/6geEVST, 2020, [Online, accessed on 07/10/2020].

[15] S. Keach, "GOING VIRAL Celebs, politicians and influencers 'spreading 5G coronavirus fake news online' - including Woody Harrelson and MIA," The Sun, https://cutt.ly/ZgeDANf, Tech. Rep., 2020, [Online, accessed on 07/10/2020].

[16] N. O'Neill, "Celebs are 'super-spreaders' of coronavirus fake news, study says," New York Post, https://nypost.com/2020/04/24/celebrities-are-super-spreaders-of-fake-coronavirus-news-study/, Tech. Rep., 2020, [Online, accessed on 07/10/2020].

[17] I. Togoh, "After Hydroxychloroquine, Trump Is Now Seeking To Get Another Unproven Drug Approved By The FDA: Report," Forbes, https://cutt.ly/bgeRSTF, Tech. Rep., 2020, [Online, accessed on 07/10/2020].

[18] BBC News, "Coronavirus: Hydroxychloroquine ineffective says Fauci," BBC News, https://www.bbc.com/news/world-us-canada-53575964, Tech. Rep., 2020, [Online, accessed on 07/10/2020].

[19] W. Y. Wang, ""Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 2017, pp. 422–426.

[20] Paula Fraga-Lamas and Tiago M. Fernández-Caramés, "Leveraging Distributed Ledger Technologies and Blockchain to Combat Fake News," *CoRR*, vol. abs/1904.05386, 2019. [Online]. Available: http://arxiv.org/abs/1904.05386

[21] 20BITS, "Cómo combatir las fake news del coronavirus (20 minutos)," https://cutt.ly/6tdtXOW, 2020, [Online, accessed on 11/03/2020].

[22] Snopes, "Snopes," https://www.snopes.com/, 1994, [Online, accessed on 19/03/2020].

[23] Politifact, "Politifact. The Poynter Institute," https://www.politifact.com/, 2017, [Online, accessed on 12/03/2020].

[24] Newtral, "Newtral," https://www.newtral.es/, 2018, [Online, accessed on 11/03/2020].

[25] Maldita.es, "Maldita.es," https://maldita.es/, 2018, [Online, accessed on 11/03/2020].

[26] RTVE, " La estructura de la noticia," https://cutt.ly/FttzBbS, 2010, [Online, accessed on 06/03/2020].

[27] A. Mosseri, "Working to Stop Misinformation and False News," https://cutt.ly/2ujbviY, Facebook, Tech. Rep., 2017, [Online, accessed on 01/03/2020].

[28] C. K. Hiramath and G. C. Deshpande, "Fake News Detection Using Deep Learning Techniques," *2019 1st International Conference on Advances in Information Technology (ICAIT)*, pp. 411–415, 2019.

[29] WhatsApp, "Consejos para prevenir la propagación de rumores y noticias falsas," https://cutt.ly/8ujx7K8, 2020, [Online, accessed on 01/03/2020].

[30] Á. Figueira and L. Olivera, "The current state of fake news: challenges and opportunities," in *CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies, CENTERIS / ProjMAN*, vol. 121, 2017, pp. 817–825.

[31] J. Y. Khan, M. T. I. Khondaker, A. Iqbal, and S. Afroz, "A Benchmark Study on Machine Learning Methods for Fake News Detection," *arXiv (arXiv:1905.04749v1)*, 2019.

[32] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, "Automatic Detection of Fake News," in *27th International Conference on Computational Linguistics*, 01 2019, pp. 3391–3401.

[33] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," in *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, I. Traore, I. Woungang, and A. Awad, Eds. Cham: Springer International Publishing, 2017, pp. 127–138.

[34] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," in *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2017, pp. 900–903.

[35] X. Zhou, R. Zafarani, K. Shu, and H. Liu, "Fake News: Fundamental Theories, Detection Strategies and Challenges," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, ser. WSDM '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 836–837. [Online]. Available: https://doi.org/10.1145/3289600.3291382

[36] N. J. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," *Proceedings of the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1–4, 2015. [Online]. Available: https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/pra2.2015.145052010082

[37] N. O'Brien, "Machine Learning for Detection of Fake News," [Online, accessed on 17/02/2021]. [Online]. Available: https://tinyurl.com/t4ia0s5m

[38] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake News Detection Using Machine Learning Ensemble Methods," vol. 2020.

[39] S. Kaur, P. Kumar, and P. Kumaraguru, "Automating fake news detection system using multi-level voting model," vol. 24, no. 12.

[40] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," Bitcoin, Bitcoin: A Peer-to-Peer Electronic Cash System, Tech. Rep. https://bitcoin.org/bitcoin.pdf, 2009, [Online, accessed on 09/03/2020]. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[41] D. Arroyo-Guardeño, J. Díaz-Vico, and L. Hernández-Encinas, *¿Qué sabemos de? Blockchain*. CSIC, 2019.

[42] P. Dunphy and F. A. P. Petitcolas, "A First Look at Identity Management Schemes on the Blockchain," *IEEE Security Privacy*, vol. 16, no. 4, pp. 20–29, 2018.

[43] C. Alonso, "Combatir Fake News con Blockchain & Ethereum," https://tinyurl.com/yd4ezj4g, 2020, [Online, accessed on 26/05/2020].

[44] O. Jacobovitz, "Blockchain for Identity Management," The Lynne and William Frankel Center for Computer Science Department of Computer Science, Ben-Gurion University, Beer Sheva, Israel, Tech. Rep., 2016.

[45] P. Dunphy, L. Garatt, and F. Petitcolas, "Decentralizing Digital Identity: Open Challenges for Distributed Ledgers," in *IEEE European Symposium on Security and Privacy Workshops*, 2017.

[46] D. Augot, H. Chabanne, O. Clémot, and W. George, "Transforming face-to-fase identity proofing into anonymous digital identity using the Bitcoin blockchain," in *15th Annual Conference on Privacy, Security and Trust*, 2017.

[47] M. Aldwairi and A. Alwahedi, "Detecting Fake News in Social Media Networks," *Procedia Computer Science*, vol. 141, pp. 215–222, 2018, the 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2018) / The 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2018) / Affiliated Workshops. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050918318210

[48] F. Dounis, "Detecting Fake News With Python And Machine Learning," The Startup, Tech. Rep., [Online, accessed on 17/02/2021]. [Online]. Available: https://tinyurl.com/110ibg6e

[49] Pantech, "Fake News Detection using Machine Learning," Pantech ProLabs India Pvt Ltd., Tech. Rep., [Online accessed on 17/02/2021]. [Online]. Available: https://tinyurl.com/yaz2xfbs

[50] J. Camacho, A. Pérez-Villegas, P. García-Teodoro, and G. Maciá-Fernández, "PCA-based Multivariate Statistical Network Monitoring for Anomaly Detection," *Computers & Security*, vol. 59, pp. 118–137, June 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167404816300116