




# Desinformación y Alfabetización Mediática: Clasificación de Patrones de Confianza en Noticias entre Adolescentes mediante Machine Learning

## Misinformation and Media Literacy: Classification of News Trust Patterns among Adolescents using Machine Learning

Marta-Lilia Eraña-Díaz<sup>1\*</sup> , Jorge Pablo Oseguera Gamba<sup>1</sup>  y Nadia Lara Ruiz<sup>2</sup> 

<sup>1</sup>Centro de Investigación en Ciencias Cognitivas (CINCCO), Universidad Autónoma del Estado de Morelos  
Av. Universidad No. 1001, Chamilpa, 62210 Cuernavaca, Mor.

<sup>2</sup>Facultad de Ciencias Químicas e Ingeniería (FCQEI), Universidad Autónoma del Estado de Morelos  
Av. Universidad No. 1001, Chamilpa, 62210 Cuernavaca, Mor  
[merana@uaem.mx](mailto:merana@uaem.mx)

### PALABRAS CLAVE:

Desinformación,  
aprendizaje  
automático,  
noticias falsas,  
confianza en los  
medios,  
cognición  
artificial,  
reproducibilidad

### RESUMEN

Este estudio examina la detección de noticias falsas entre adolescentes utilizando técnicas de Machine Learning. Se plantean dos preguntas fundamentales: cómo reaccionan los jóvenes ante la desinformación y qué nivel de confianza tienen en la veracidad de las noticias. A través de un análisis detallado con algoritmos de aprendizaje automático, como k-means, Decision Trees y Random Forest, se clasificaron las respuestas de los participantes, revelando patrones significativos en su confianza hacia los medios de comunicación. Los hallazgos sobre esta muestra indican que aproximadamente el 15% de los adolescentes no confían plenamente en las noticias y casi el 50% carece de habilidades para identificar información falsa. Esto resalta la necesidad urgente de implementar programas de alfabetización mediática que fortalezcan la capacidad crítica de los jóvenes para discernir la veracidad de la información en un entorno digital saturado de desinformación. Además, el estudio sugiere que futuras investigaciones deben explorar otros algoritmos de Machine Learning y evaluar el impacto de intervenciones educativas en la alfabetización mediática.

### KEYWORDS:

Misinformation,  
machine  
learning, fake  
news, media  
trust, artificial  
cognition,  
reproducibility

### ABSTRACT

This study examines the detection of fake news among adolescents using Machine Learning techniques. It poses two fundamental questions: how do young people react to misinformation, and what level of trust do they have in the truthfulness of news. Through a detailed analysis employing Machine Learning algorithms such as k-means, Decision Trees, and Random Forest, participants' responses were classified, revealing significant patterns in their trust towards media. The findings in this study, indicate that approximately 15% of adolescents do not fully trust the news, and almost 50% lack the skills to identify false information. This underscores the urgent need to implement media literacy programs that enhance young people's critical capacity to discern the truth of information in a digital environment saturated with misinformation. Additionally, the study suggests that future research should explore other Machine Learning algorithms and assess the impact of educational interventions on media literacy.

• Recibido: 28 de noviembre de 2024 • Aceptado: 11 de marzo de 2025 • Publicado en línea: 10 de febrero de 2026

## 1. INTRODUCCIÓN

En la era digital actual, la proliferación de noticias falsas representa un desafío significativo para la sociedad, especialmente entre los jóvenes que son ávidos consumidores de información en línea. La capacidad de discernir la veracidad de las noticias se ha vuelto crucial en un entorno donde la desinformación puede propagarse rápidamente a través de las redes sociales y otras plataformas en línea. Según reporte reciente de Reuters [1] a nivel mundial sólo 40% de los encuestados consideran que se puede confiar en la mayoría de las noticias. En el caso de México, el porcentaje es de 35%, y en el caso de los encuestados de 18 a 24 años baja a 28%. A pesar de la creencia generalizada de que los adolescentes son expertos en la detección de noticias falsas, investigaciones recientes han demostrado que su habilidad en este aspecto puede ser limitada.

La desinformación en la web social es un fenómeno que implica la difusión de información falsa o engañosa a través de plataformas de redes sociales. Una de las características de la desinformación es la rapidez de difusión, ya que las redes sociales permiten que la información, verdadera o falsa, se difunda rápidamente y a gran escala debido a la naturaleza viral del contenido. El uso de algoritmos de recomendación de plataformas como Facebook, Twitter, YouTube, entre otras, tienden a priorizar el contenido que genera más interacción, lo que puede incluir información sensacionalista o falsa, además la facilidad de creación de contenido con la inteligencia artificial generativa permite que cualquiera puede crear y compartir contenido en las redes sociales, facilitando la proliferación de información no verificada. Especialmente preocupante fue el descubrimiento de que las noticias falsas se esparcen más rápido que las verdaderas [2].

El presente estudio se enfoca en analizar la percepción de los adolescentes sobre las

noticias falsas y su capacidad para identificar información engañosa en línea. A través de la aplicación de técnicas de *Machine Learning*, se busca profundizar en la confianza a los medios de comunicación y detectar las acciones frente a las noticias falsas entre esta población. Al comprender mejor cómo los adolescentes interactúan con la información en línea y cómo pueden mejorar sus habilidades de discernimiento, se pretende contribuir al desarrollo de estrategias educativas y preventivas para combatir la desinformación en la web y promover la alfabetización mediática entre los jóvenes.

Por otra parte, la inteligencia artificial (IA) ha emergido como una herramienta transformadora, capaz de abordar una amplia gama de problemas complejos en diversas disciplinas. En este contexto, la cognición artificial se refiere al conjunto de procesos y capacidades que permiten a los sistemas de IA emular, simular o replicar funciones cognitivas humanas, como el aprendizaje, la percepción, el razonamiento, la toma de decisiones y la comprensión del lenguaje. A través de algoritmos y modelos computacionales, la cognición artificial busca imitar la forma en que los seres humanos procesan información, interpretan datos y responden a su entorno, permitiendo a las máquinas realizar tareas que tradicionalmente requieren inteligencia humana. El aprendizaje automático (ML), una subdisciplina de la IA, ha demostrado su eficacia en la identificación de patrones y la toma de decisiones basadas en datos. Sin embargo, su aplicación no es tan sencilla como podría parecer. La creación de modelos de ML efectivos requiere un enfoque cuidadoso y metódico, ya que la generación de "malos modelos" puede resultar en sesgos y en la omisión de casos minoritarios, lo que puede tener consecuencias significativas en la interpretación de los resultados. Por lo tanto, es crucial que los desarrolladores y científicos de datos comprendan las pautas necesarias para construir modelos robustos y justos. Un aspecto fundamental es la reproducibilidad en el aprendizaje automático

[3], [4] siendo esta la capacidad de regenerar un modelo con precisión, garantizando que las inferencias sean consistentes y transparentes. Sin embargo, la naturaleza inherente de los números pseudoaleatorios utilizados en la generación de modelos puede complicar este proceso.

Las técnicas de aprendizaje no supervisado realizan clasificaciones sin disponer de etiquetas en los datos, lo que implica que las categorías son desconocidas de antemano [5]. En el caso de los datos continuos, el objetivo del análisis de agrupamiento (clustering) es organizar los elementos en grupos distintos, de manera que los elementos dentro de un mismo grupo sean lo más similares posible entre sí y lo más diferentes posible de los elementos en otros grupos. Dicho de otro modo, los *clusters* deben mostrar diferencias claras entre ellos, mientras que los elementos dentro de cada *cluster* deben ser altamente similares. En cuanto a las técnicas de aprendizaje supervisado, se destacan los árboles de decisión, incluyendo los del tipo CART (*Classification and Regression Trees*). La creación de un árbol de decisión es un proceso iterativo que comienza ubicando toda la muestra disponible en el nodo raíz. A partir de ahí, se realizan sucesivas particiones que generan las ramas del árbol, hasta llegar a los nodos terminales o hojas, los cuales están formados por conjuntos de individuos que han respondido de manera similar a ciertos ítems. Para evitar que el árbol resultante sea poco fiable debido a las particiones, se utiliza una técnica de validación como criterio de parada. Esto se logra dividiendo la muestra en dos partes: una de entrenamiento (70 u 80%) y otra de validación (30 o 20%). Para un mayor entendimiento de estos temas, se recomienda consultar [6], [7].

En el presente trabajo se aplicaron técnicas de ML para el tratamiento de los datos obtenidos de los resultados del Proyecto "Teen Fake News Poll on After School" [8]. Los algoritmos y scripts de ML, se realizaron en

Python que es un lenguaje libre altamente utilizado en la ciencia de datos dada la sencillez y claridad de su sintaxis [9]. El resto del documento está organizado como sigue: la Sección 2 presenta trabajos relacionados recientes y relevantes en la investigación de la aplicación de inteligencia artificial. La Sección 3 describe el método realizado para estudiar la desinformación con respecto a las noticias falsas, FN. En la Sección 4 se presenta el análisis y los resultados, por último, en la sección 5, la discusión y las conclusiones, así como también los trabajos a futuro.

## 2. ESTUDIOS RELACIONADOS

Las noticias falsas son un fenómeno que es importante investigar, pues son causantes de polarización, inestabilidad social y daños a la democracia [10]. Varios estudios han aplicado machine learning a la detección de noticias falsas [11], sin embargo, en este estudio la aplicación de esta tecnología servirá no para detectar noticias falsas, sino para entender mejor el fenómeno. Existen diversos factores involucrados en la creencia de noticias falsas [12], [13] los cuales se pueden dividir en tres grupos. Uno es la característica del mensaje, por ejemplo, qué tan consistente es el mensaje con el resto de las creencias del receptor. El segundo grupo consiste en características del receptor, por ejemplo, sus estilos cognitivos, sus predisposiciones y su grado de alfabetización mediática. Por último, se encuentran factores contextuales, como si hay alguna alerta por parte del medio sobre posible información no verificada o si hay algunas de las advertencias que algunas redes sociales hacen sobre cierto contenido. En el caso específico de los adolescentes, existen estudios sobre qué factores los llevan a consumir y compartir noticias falsas [14].

Para destacar cómo la aplicación de técnicas de machine learning han sido beneficiosas en diferentes campos [15] y cómo estos enfoques pueden ser relevantes para el estudio de detección de noticias falsas en adolescente se revisó un estudio realizado por Smith et al.

[16] donde se presenta una metodología para automatizar la detección de narrativas, redes y actores influyentes de desinformación, integrando procesamiento de lenguaje natural y aprendizaje automático para cuantificar el impacto en la difusión de noticias falsas. En Langley et al. [17] se construyen y analizan varios modelos de *machine learning* con bajos niveles de error para garantizar la selección del modelo óptimo para determinar cómo las noticias falsas han afectado a la sociedad y el impacto en el futuro de las sociedades. La investigación de Verma et al. [18] se centra en una revisión sistemática de las tecnologías de detección de noticias falsas existentes mediante la exploración de diversas técnicas de aprendizaje automático antes y después de la pandemia de COVID-19.

La aportación de este estudio es generar un método que amplíe los resultados de respuestas sobre desinformación, en especial de las noticias falsas (FN) aplicando técnicas de inteligencia artificial (IA). Por lo tanto, los objetivos de este estudio fueron:

- Realizar un análisis con *Machine Learning* (ML) con el fin de determinar los atributos diferenciadores de las respuestas de los participantes en los diferentes lugares donde se recolectó la muestra, sobre Confianza en los medios (*Trust-Media*, TM).
- Probar la importancia de la variabilidad en el entrenamiento del modelo para obtener resultados sólidos y confiables.

### 3. ESTUDIO SOBRE FAKE NEWS PARA EVALUAR EL NIVEL DE CONFIANZA EN LOS MEDIOS

El problema central de este estudio es aplicar técnicas de *machine learning* para determinar qué factores influyen en la confianza de los individuos hacia los medios de comunicación ya que la creciente

desconfianza hacia los medios puede llevar a la desinformación y a la apatía cívica.

Para conocer la opinión de adolescentes se implementó en la aplicación *After School* [19] una votación (*Poll*) de cinco preguntas sobre FN. *After School* es una plataforma en línea para administrar los datos asociados con clases de tutoría, es una aplicación fácil de usar con características como asistencia en línea, envío de tareas, informes detallados de rendimiento y mucho más, es una app muy bien aceptada por los estudiantes, padres y tutores. Esta herramienta fue utilizada en el proyecto *Teen Fake News Poll* con la cual se encuestó a sus usuarios sobre el tema de las FN como se mencionó en el párrafo anterior. Durante varios días más de 39,000 estudiantes de secundaria en los 50 estados de la unión americana del norte, participaron en la encuesta. Las preguntas que se realizaron fueron:

- 1 ¿Has oído hablar de las “noticias falsas”?
- 2 ¿Eres bueno detectando noticias falsas?
- 3 ¿Crees que la mayoría de los artículos periodísticos son ciertos?
- 4 Si ves a alguien compartiendo noticias falsas, ¿qué haces?
- 5 Si un artículo de noticias no dijera la verdad, ¿crees que podrías saberlo?

Los datos completos pueden ser consultados en la página del proyecto a través de *data.world* [20] que es un almacén electrónico donde generalmente empresas u organización pueden mantener una gran cantidad de información, en este datawarehouse el almacenaje es seguro, fiable, fácil de recuperar y fácil de administrar. En las siguientes secciones se presenta más información sobre este dataset.

#### 3.1 Metodología

La metodología que describe, en forma general el proceso que se llevó a cabo para la realización de este estudio de los niveles de Trust-media a partir de los resultados del

proyecto *Teen Fake News Poll* realizado con la app *After School*, se presenta en la Figura 1.

### 3.1 Participantes y Recolección de Datos

Se utilizó un dataset recopilado en 2017 que contenía las respuestas de adolescentes a una encuesta sobre FN.

Los resultados de la encuesta se muestran en tres bases de datos: Por pregunta, por ciudad y por estado. Más de 39 mil adolescentes contestaron la encuesta (Figura 2). Las características de las bases de datos del proyecto se presentan en el apéndice A donde se especifica el detalle de este proyecto y los resultados publicados.

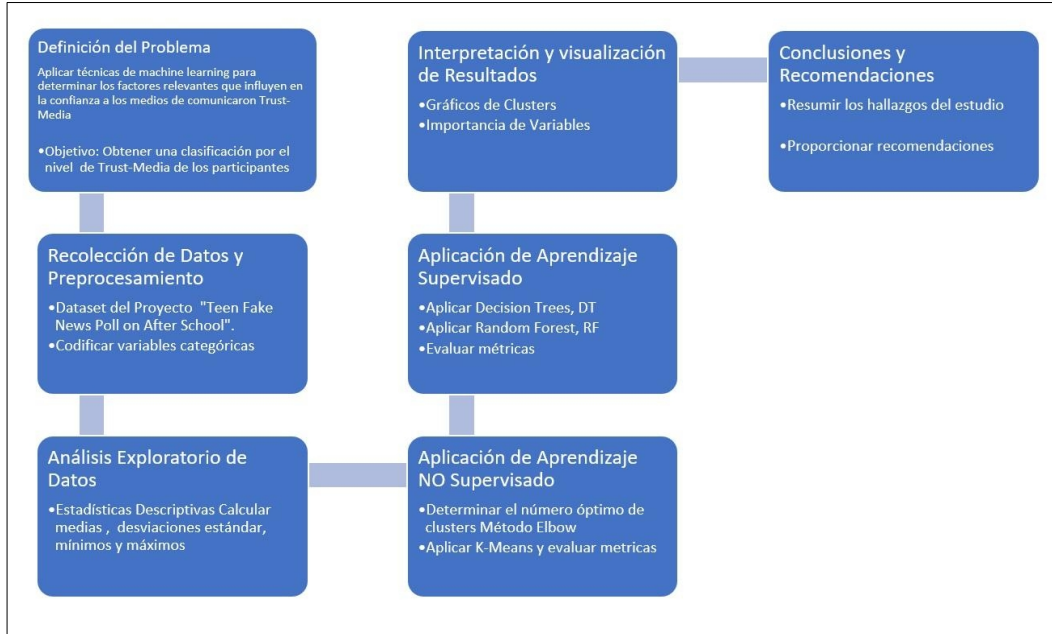


Figura 1. Diagrama de flujo de los pasos seguidos en este estudio.

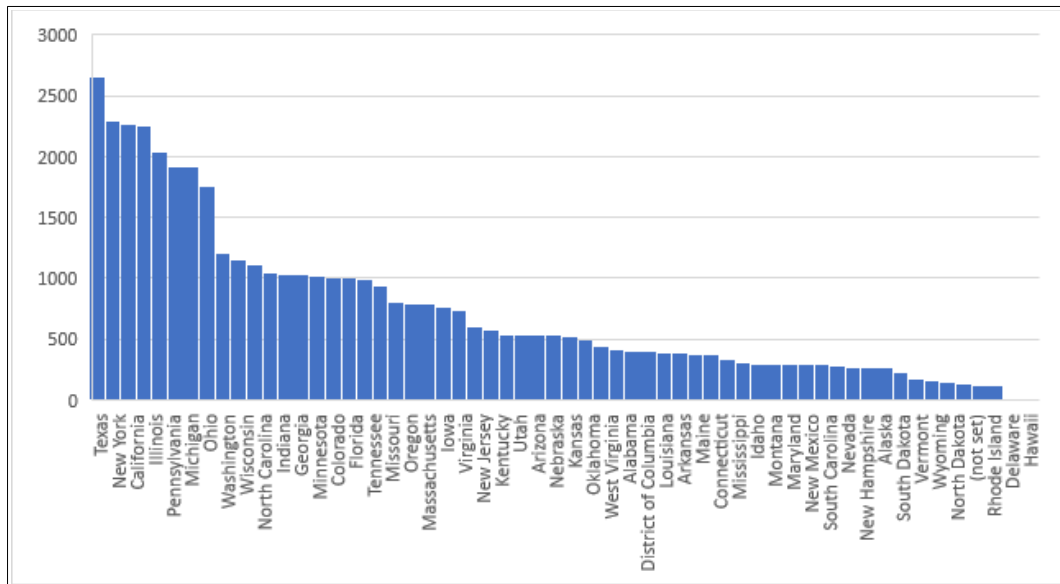


Figura 2. Número de participantes por Estado incluidos 132 que no se identificaron (not set), en total fueron 39,005.

### 3.2 Materiales y Recursos Tecnológicos

1) **Algoritmos de ML:** La selección de los algoritmos de machine learning se basó en el conocimiento de su desempeño [21], [22]. Para el aprendizaje no supervisado, se utilizó k-means, mientras que para el aprendizaje supervisado se emplearon árboles de decisión (DT) y bosques aleatorios (RF). Es interesante ilustrar algunos procesos cognitivos con estos algoritmos de aprendizaje artificial para facilitar la comprensión del funcionamiento de estos y la capacidad para tomar decisiones basadas en datos, esto resalta también la flexibilidad y diversidad de los resultados posibles. Podemos comparar el proceso de clasificación en un árbol de decisión con la entrada de personas a un espacio donde los objetos están ubicados de manera específica. Si cada persona ingresa por una puerta diferente, comenzará a clasificar el objeto más cercano a ella, tomando decisiones basadas en características visibles. Al final, podríamos tener tantas clasificaciones distintas como puertas de entrada, multiplicado por el

número de personas que ingresaron, lo que refleja la diversidad de enfoques y resultados posibles. Este proceso de toma de decisiones se asemeja a cómo un árbol de decisión segmenta los datos en función de características específicas, creando ramas que representan diferentes decisiones y resultados. En el caso del algoritmo K-means, la analogía puede ser vista en cómo un grupo de personas se agrupa en un espacio. Imaginemos que un grupo de personas se encuentra en un parque y, al observar el entorno, comienzan a agruparse según sus preferencias, como la cercanía a un árbol, una fuente o un banco. Cada persona se une al grupo más cercano, formando clústeres basados en similitudes. De manera similar, K-means agrupa datos en clústeres, donde cada punto de datos se asigna al clúster más cercano, optimizando la cohesión dentro de cada grupo.

2) **Base de datos de los resultados del proyecto en los porcentajes de respuestas por estado a cada una de las preguntas de la encuesta:** La base de datos tiene 51 renglones y 11 columnas (ver apéndice A). Las preguntas y opciones de respuesta están presentadas en la tabla 1.

Tabla 1. Preguntas y respuestas del cuestionario.

Q1		Q2		Q3		Q4			Q5	
Have you heard of "fake news stories?"		Are you good at spotting fake news stories?		Do you believe most news articles are true		If you see someone sharing fake news stories, what do you do?			If a news article wasn't telling the truth, do you think you could tell?	
1a	1b	2a	2b	3a	3b	4a	4b	4c Share	5a	5b
Yes, I have heard of fake news	No, I have not heard of fake news	Yes, I'm good at spotting fake news	No, I'm bad at spotting fake news	Yes, I believe most news articles are true	No, I believe most news articles are false	Ignore it	Call them out	it if you like it	Yes, think I could tell if an article was telling the truth"	No, I don't think I could tell if an article was telling the truth

3) **Software y paquetería de Python:** Para implementar estos algoritmos, utilizamos software y paquetes de Python. Python ofrece varias ventajas para el desarrollo y aplicación de algoritmos de machine learning por la facilidad de uso y simplicidad, hay una amplia comunidad y soporte contando con una extensa

comunidad de desarrolladores y una gran cantidad de recursos disponibles, incluyendo tutoriales, documentación y foros de discusión. Hay numerosas bibliotecas y paquetes en Python específicamente diseñados para machine learning, en este estudio se utilizó scikit-learn, NumPy, Pandas [23]. Sin embargo,

también hay algunas limitaciones al utilizar Python para machine learning ya que es un lenguaje interpretado, lo que significa que puede ser más lento en comparación con lenguajes compilados como C++ o Java. Con respecto a la gestión de memoria, las librerías NumPy y pandas optimizan la gestión de esta sin embargo para gran número de datos, Python puede no ser la mejor opción debido a su consumo de memoria. Por último, aunque las librerías de Python para machine learning son potentes, su uso a veces puede limitarse a las funcionalidades que estas librerías proporcionan. Personalizar o extender estas herramientas más allá de su diseño inicial puede ser complicado y requiere un conocimiento profundo del código subyacente.

- 4) **Recursos Computacionales:** Se utilizó una laptop HP equipada con un procesador 12th Gen Intel(R) Core(TM) i7-1255U a 1.70 GHz con 16 GB de RAM, y sistema operativo de 64 bits, procesador x64, Windows 11. Se desarrolló en la plataforma Jupyter Notebook, que proporciona un entorno interactivo ideal para el desarrollo y la ejecución de scripts en Python.

### 3.3. Procedimiento

Como se muestra en la Figura 1, para analizar los datos del dataset y evaluar la capacidad de los adolescentes para detectar noticias falsas, se aplicaron diversas técnicas de aprendizaje automático. En primer lugar, se utilizó el aprendizaje no supervisado mediante *k-means* para identificar patrones y agrupamientos en los datos. Posteriormente, se emplearon técnicas de aprendizaje supervisado, incluyendo *Decision Trees* (DT) y *Random Forest* (RF), para desarrollar modelos predictivos. Es crucial destacar la importancia de correr el modelo con varias semillas para la generación de números pseudoaleatorios. Al utilizar la librería random en Python para inicializar el generador de números

aleatorios, se introduce variabilidad en el proceso de entrenamiento del modelo, lo que permite explorar diferentes configuraciones y mejorar la robustez de los resultados obtenidos. Además, la importancia de la reproducibilidad en el aprendizaje automático (ML) permite generar un modelo de ML que garantice precisión y transparencia, así una forma de garantizar que los modelos sean confiables es gestionar los números aleatorios producidos durante el entrenamiento del modelo [24], [4]. En el apéndice B se presenta un marco teórico sobre los algoritmos utilizados en el presente estudio.

## 4. ANÁLISIS Y RESULTADOS

### 4.1. Aprendizaje No Supervisado

El algoritmo aplicado es *k-means* y mediante el método de Elbow [25] se determinó el mejor número de clústeres para realizar el agrupamiento. Se aplicó dos veces el método obteniendo agrupaciones similares.

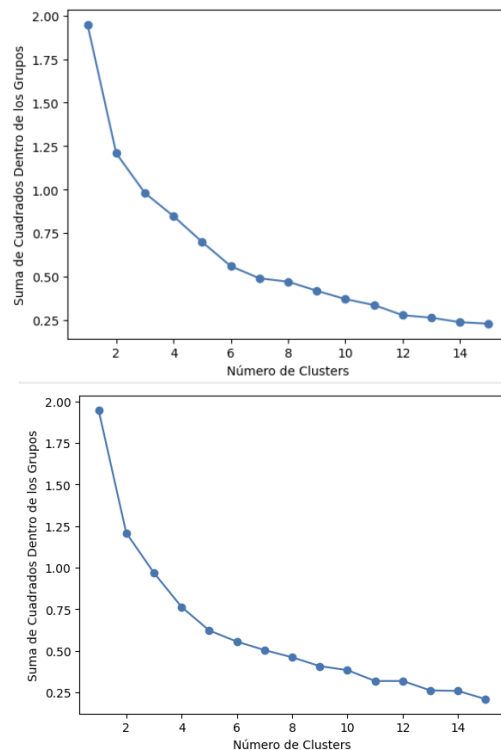


Figura 3. Gráficas del Método Elbow con dos semillas aleatorias diferentes.

El método de Elbow [25] muestra el número de clústeres en relación con la disminución de la Suma de Cuadrados, SC con el algoritmo *k-means*. En ambos casos una agrupación con  $k=4$  es buena, aunque no está bien marcada la velocidad de disminución con la integración de más clústeres, pero consideramos adecuada para realizar este análisis.

En la Tabla 2 y 3 se presenta la efectividad del agrupamiento realizado por *K-means* de ambos modelos. La distancia promedio dentro del grupo, es la distancia promedio de todos los puntos dentro de un mismo clúster a su centroide (punto central del clúster), cuanto menor sea esta distancia, mejor será la cohesión del clúster. Una baja distancia intra-cluster indica que los puntos dentro del mismo clúster están más cerca unos de otros y, por lo tanto, el clúster es más compacto. Por otra parte, un alto porcentaje de  $\text{between\_SS} / \text{total\_SS}$  indica que los clusters están bien separados entre sí, es decir, las diferencias entre los clusters son significativas. Estas métricas ayudan a evaluar qué tan bien el algoritmo de agrupación separa los puntos de datos en grupos significativos en función de similitudes o diferencias en sus atributos. El mejor desempeño es el del modelo UNO.

**Tabla 2.** Métricas del Agrupamiento. Modelo UNO.

Cluster	% tamaño cluster	Suma de cuadrados interna por grupo	Distancia promedio Intra-cluster	Porcentaje de $\text{between\_SS} / \text{total\_SS}$
1	24	32		
2	2	11	0.1142	60.9724
3	14	14		
4	11	13		

**Tabla 3.** Métricas del Agrupamiento Modelo DOS.

Cluster	% tamaño cluster	Suma de cuadrados interna por grupo	Distancia promedio Intra-cluster	Porcentaje de $\text{between\_SS} / \text{total\_SS}$
1	12	47		
2	25	18	0.1164	58.3723
3	12	11		
4	2	14		

En la siguiente tabla se muestran los 50 estados en el agrupamiento realizado por *k-means* en cuatro clusters del Modelo uno.

**Tabla 4.** Agrupamiento realizado en cuatro clústeres por el modelo UNO.

Clúster 1	Clúster 2	Clúster 3	Clúster 4
Alaska	Hawaii	Arkansas	Alabama
Arizona	Rhode Island	Delaware	Florida
California		Georgia	Nebraska
Colorado		Idaho	New
Connecticut		Indiana	Hampshire
District of Columbia		Iowa	North Dakota
Illinois		Kentucky	Tennessee
Kansas		Maine	Utah
Louisiana		Minnesota	Vermont
Maryland		Ohio	Washington
Massachusetts		Pennsylvania	West Virginia
Michigan		South Dakota	Wyoming
Mississippi		Texas	
Missouri		Wisconsin	
Montana			
Nevada			
New Jersey			
New Mexico			
New York			
North Carolina			
Oklahoma			
Oregon			
South Carolina			
Virginia			

En todos los estados la mitad o más de la mitad de los respondientes contestaron lo siguiente:

- 1a: Sí, he escuchado sobre “noticias falsas” (.77-.88)
- 2a: Sí, soy bueno identificando noticias falsas (.57-.8)
- 4b: Si veo a alguien compartiendo historias de noticias falsas, llamo la atención al respecto. (.5-.69)

Como se muestra en la Tabla 5, las respuestas variaron en las preguntas 3 (¿Crees que la mayoría de los artículos noticiosos son verdaderos?) y 5 (Si un artículo noticioso no dijera la verdad, ¿podrías notarlo?). Son estas variaciones las que los clusters identifican.

Al analizar las diferencias, notamos que surgen tres perfiles de confianza en los clusters, como se muestra en la Tabla 6. A la pregunta 3, la mitad o más contestaron creer

que la mayoría de los artículos noticiosos son falsos, con excepción del clúster 2: En Hawaii y en Rhode Island la mayoría contestó creer que la mayoría de los artículos noticiosos son verdaderos, así como creer poder identificar que si un artículo noticioso no dijera la verdad. En otras palabras, lo que caracteriza al clúster 2 es que confían tanto en las noticias como en su capacidad de identificar las noticias falsas. Este es el perfil más vulnerable para caer en noticias falsas, pues al creer que la mayoría de las noticias son verdaderas, no se mantienen alerta. Más aún creen que pueden identificarlas fácilmente, lo que puede llevar a un exceso de confianza.

**Tabla 5.** Perfiles de los clústers.

	Q1	Q2	Q3	Q4	Q5
Cluster 1	1a	2a	3b	4b	<b>0.45-0.6</b>
Cluster 2	1a	2a	<b>3a</b>	4b	5i
Cluster 3	1a	2a	3b	4b	<b>5b</b>
Cluster 4	1a	2a	3b	4b	5a

En contraste, en los estados del clúster 3 sucede lo opuesto, la mayoría desconfía tanto en las noticias como en su capacidad de identificar las noticias falsas. Este es el perfil menos vulnerable a las noticias falsas, pues la desconfianza los lleva mantenerse alertas de las noticias falsas. Sin embargo, este perfil también tiene sus problemas. El desconfiar tanto puede llevar a no confiar en nada en lo absoluto y llevar a apatía y dejar de informarse.

En el caso del clúster 4, desconfían en las noticias pero confían en su capacidad de identificar las falsas. Este podría ser el perfil óptimo, pues se desconfía de las noticias, lo que los mantiene alerta, pero al confiar en su capacidad de detectar noticias falsas, las analizan con cuidado. En los estados del clúster 1, el más grande, la mayoría desconfían en las noticias, pero no hay una tendencia clara sobre su confianza en la capacidad de identificar las noticias falsas; van del 45% al 60% en la respuesta 5b (desconfianza en su capacidad de identificar noticias falsas). Por lo tanto, algunos del

clúster 1 tienen el perfil del clúster 3 y otros del 4.

**Tabla 6.** Perfiles de confianza en los clústers.

	Confía en las noticias	Confía en su capacidad de identificar noticias falsas
Cluster 1	No	Si/No
Cluster 2	Si	Si
Cluster 3	No	No
Cluster 4	No	Si

## 4.2. Aprendizaje Supervisado

Con la base de datos ampliada con el número del clúster asignado en el aprendizaje No supervisado, se aplicaron los algoritmos de DT y RF.

Se ejecutaron múltiples pruebas utilizando diferentes semillas, obteniendo las métricas correspondientes al agrupamiento del Modelo UNO. La muestra se dividió en un 70% de los datos para el entrenamiento y un 30% para la prueba. Esta partición se realizó de manera aleatoria, ya que se sabe que la forma en que se divide el conjunto de datos puede influir en los resultados. Por ello, se llevaron a cabo varias ejecuciones con distintas semillas. En la tabla 7 se presentan las métricas obtenidas donde el promedio de la precisión es 0.825 con una desviación estándar de 0.112 para DT y 0.873 de promedio con 0.081 de desviación estándar para RF.

**Tabla 6.** Métricas con las diferentes semillas del Modelo UNO (sobre test data).

Semilla aleatoria	Precisión del DT	Precisión del RF
384378	0.75	0.818182
627516	0.75	0.818182
474857	0.875	0.909091
337390	0.75	0.818182
430453	1.0	1.0

El desempeño del modelo de DT es del 100% con la semilla 430453. Se observa que la matriz de confusión solo tiene 3 grupos ya que el clúster 2 solo tiene dos elementos y en la separación del dataset en la prueba (test data) no se incluyeron. La Matriz de Confusión es  $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ .

La Fig. 4 se interpreta de la siguiente manera:

```

Si 5a < 0.575 entonces
  si 4a < 0.315 entonces
    cluster 4
  otro
    si 3a <= 0.495 entonces
      si 5b < 0.45
        cluster 1
      otro
        cluster 3
    otro
      cluster 2
  otro
    cluster 1
    
```

Al aplicar RF, se obtienen las variables más relevantes, Fig. 5, en la siguiente imagen se pueden observar estas con el número de pregunta y la respuesta a, b y en el caso de la pregunta 4, c también.

Se enlistan las preguntas que resultaron más relevantes y sus respuestas en negrillas.

Q5: *If a news article wasn't telling the truth, do you think you could tell?*

-5a Yes, think I could tell if an article was telling the truth".

-5b No, I don't think I could tell if an article was telling the truth.

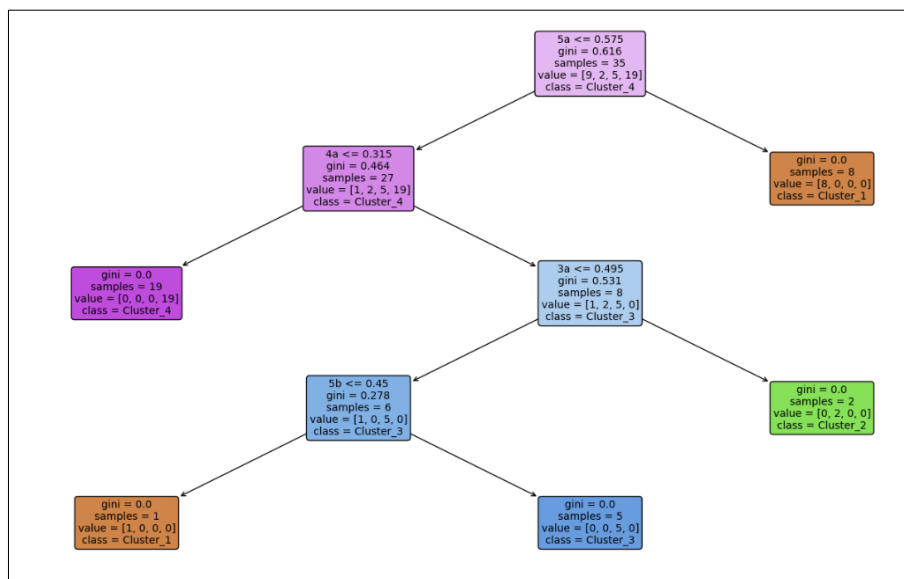


Figura 4. Árbol de Decisión reactivos seleccionados, 5a, 4a, 3a y 5b.

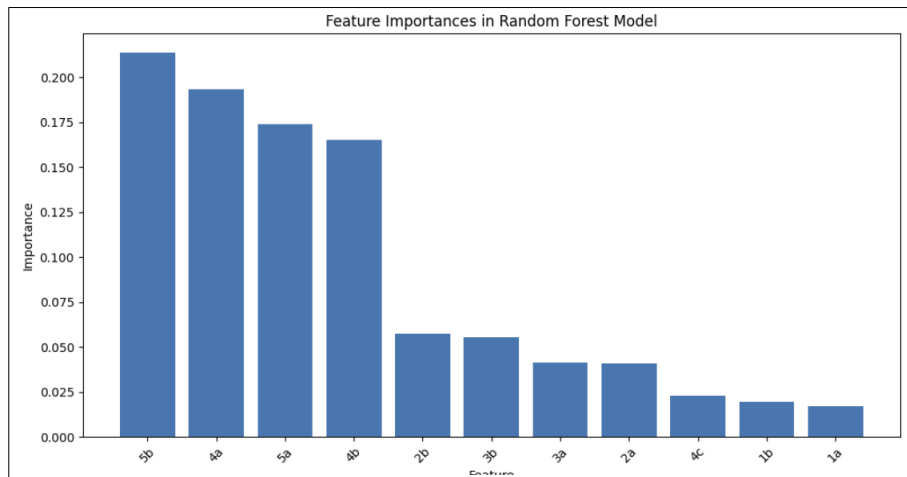


Figura 5. Variables más relevantes del RF.

Q4: *If you see someone sharing fake news stories, what do you do?*

- 4a Ignore it
- 4b Call them out
- 4c Share it if you like it

En la tabla 8 se muestran los estados con los datos y el clúster asignado donde se especifica con colores el nivel de confianza (TM) que presentaron los adolescentes de esas regiones del dataset.

**Tabla 8.** Datos de los estados presentados en orden alfabético con el clúster asignado.

State	1a	1b	2a	2b	3a	3b	4a	4b	4c	5a	5b	Cluster Modelo UNO
Alabama	0.85	0.15	0.71	0.29	0.31	0.69	0.3	0.64	0.06	0.61	0.39	Verde Oscuro
Alaska	0.8	0.2	0.6	0.4	0.42	0.58	0.28	0.64	0.08	0.4	0.6	Verde Claro
Arizona	0.79	0.21	0.64	0.36	0.27	0.73	0.32	0.62	0.06	0.46	0.54	Verde Claro
Arkansas	0.79	0.21	0.67	0.33	0.39	0.61	0.32	0.59	0.08	0.64	0.36	Verde Claro
California	0.82	0.18	0.67	0.33	0.37	0.63	0.34	0.59	0.07	0.5	0.5	Verde Claro
Colorado	0.82	0.18	0.65	0.35	0.34	0.66	0.3	0.64	0.06	0.53	0.47	Verde Claro
Connecticut	0.84	0.16	0.68	0.32	0.28	0.72	0.33	0.6	0.07	0.42	0.58	Verde Claro
Delaware	0.86	0.14	0.8	0.2	0.5	0.5	0.24	0.67	0.09	0.75	0.25	Verde Claro
District of Columbia	0.82	0.18	0.67	0.33	0.35	0.65	0.35	0.58	0.07	0.53	0.47	Verde Claro
Florida	0.84	0.16	0.72	0.28	0.27	0.73	0.36	0.58	0.07	0.58	0.42	Verde Claro
Georgia	0.82	0.18	0.65	0.35	0.37	0.63	0.33	0.59	0.07	0.6	0.4	Verde Claro
Hawaii	0.87	0.13	0.68	0.32	0.75	0.25	0.42	0.5	0.08	0.25	0.75	Rojo
Idaho	0.78	0.22	0.62	0.38	0.38	0.62	0.31	0.63	0.06	0.56	0.44	Verde Claro
Illinois	0.83	0.17	0.65	0.35	0.32	0.68	0.3	0.61	0.08	0.55	0.45	Verde Claro
Indiana	0.82	0.18	0.64	0.36	0.37	0.63	0.29	0.64	0.08	0.57	0.43	Verde Claro
Iowa	0.85	0.15	0.57	0.43	0.44	0.56	0.28	0.65	0.08	0.54	0.46	Verde Claro
Kansas	0.83	0.17	0.58	0.42	0.37	0.63	0.26	0.66	0.08	0.47	0.53	Verde Claro
Kentucky	0.87	0.13	0.6	0.4	0.39	0.61	0.3	0.61	0.09	0.61	0.39	Verde Claro
Louisiana	0.81	0.19	0.65	0.35	0.35	0.65	0.35	0.59	0.06	0.48	0.52	Verde Claro
Maine	0.83	0.17	0.64	0.36	0.34	0.66	0.25	0.69	0.06	0.56	0.44	Verde Claro
Maryland	0.85	0.15	0.68	0.32	0.23	0.78	0.4	0.52	0.08	0.5	0.5	Verde Claro
Massachusetts	0.83	0.17	0.66	0.34	0.38	0.62	0.34	0.59	0.07	0.5	0.5	Verde Claro
Michigan	0.83	0.17	0.67	0.33	0.32	0.68	0.29	0.64	0.07	0.5	0.5	Verde Claro
Minnesota	0.84	0.16	0.62	0.38	0.33	0.67	0.27	0.67	0.06	0.56	0.44	Verde Claro
Mississippi	0.82	0.18	0.65	0.35	0.31	0.69	0.31	0.65	0.05	0.53	0.47	Verde Claro
Missouri	0.8	0.2	0.64	0.36	0.33	0.67	0.29	0.63	0.07	0.53	0.47	Verde Claro
Montana	0.8	0.2	0.6	0.4	0.34	0.66	0.31	0.65	0.05	0.52	0.48	Verde Claro
Nebraska	0.83	0.17	0.57	0.43	0.3	0.7	0.31	0.61	0.08	0.57	0.43	Verde Claro
Nevada	0.8	0.2	0.66	0.34	0.35	0.65	0.29	0.64	0.07	0.52	0.48	Verde Claro
New Hampshire	0.85	0.15	0.64	0.36	0.27	0.73	0.33	0.58	0.09	0.56	0.44	Verde Claro
New Jersey	0.83	0.17	0.65	0.35	0.31	0.69	0.38	0.53	0.09	0.53	0.47	Verde Claro
New Mexico	0.85	0.15	0.68	0.32	0.37	0.63	0.36	0.56	0.08	0.45	0.55	Verde Claro
New York	0.85	0.15	0.65	0.35	0.34	0.66	0.33	0.6	0.07	0.48	0.52	Verde Claro
North Carolina	0.83	0.17	0.65	0.35	0.3	0.7	0.35	0.58	0.07	0.54	0.46	Verde Claro
North Dakota	0.84	0.16	0.6	0.4	0.2	0.8	0.23	0.69	0.09	0.5	0.5	Verde Claro
Ohio	0.84	0.16	0.64	0.36	0.41	0.59	0.3	0.63	0.07	0.55	0.45	Verde Claro
Oklahoma	0.79	0.21	0.59	0.41	0.32	0.68	0.29	0.63	0.08	0.5	0.5	Verde Claro
Oregon	0.82	0.18	0.63	0.37	0.36	0.64	0.29	0.64	0.07	0.51	0.49	Verde Claro
Pennsylvania	0.84	0.16	0.63	0.37	0.38	0.62	0.3	0.61	0.09	0.56	0.44	Verde Claro
Rhode Island	0.83	0.17	0.7	0.3	0.61	0.39	0.35	0.57	0.08	0.33	0.67	Rojo
South Carolina	0.81	0.19	0.64	0.36	0.33	0.67	0.31	0.63	0.06	0.53	0.47	Verde Claro
South Dakota	0.83	0.17	0.68	0.32	0.39	0.61	0.25	0.66	0.08	0.54	0.46	Verde Claro
Tennessee	0.82	0.18	0.63	0.37	0.31	0.69	0.32	0.6	0.08	0.6	0.4	Verde Claro
Texas	0.81	0.19	0.65	0.35	0.34	0.66	0.35	0.57	0.08	0.59	0.41	Verde Claro
Utah	0.77	0.23	0.65	0.35	0.26	0.74	0.3	0.65	0.05	0.62	0.38	Verde Claro
Vermont	0.86	0.14	0.57	0.43	0.13	0.87	0.35	0.61	0.04	0.67	0.33	Verde Claro
Virginia	0.84	0.16	0.64	0.36	0.36	0.64	0.34	0.59	0.07	0.4	0.6	Verde Claro
Washington	0.82	0.18	0.64	0.36	0.24	0.76	0.32	0.62	0.06	0.6	0.4	Verde Claro
West Virginia	0.84	0.16	0.65	0.35	0.26	0.74	0.24	0.68	0.08	0.53	0.47	Verde Claro
Wisconsin	0.85	0.15	0.64	0.36	0.47	0.53	0.25	0.69	0.06	0.55	0.45	Verde Claro
Wyoming	0.84	0.16	0.67	0.33	0.29	0.71	0.25	0.68	0.07	0.61	0.39	Verde Claro

Se observa que el clúster del color verde oscuro, son los que tienen una muy baja TM, y el clúster rojo con 2 estados, Hawaii y Rhode Island son los que más confían, le

siguen los de color amarillo y en verde claro son los que tiene baja confianza.

En las siguientes imágenes, Fig. 6 y 7, se

observan los porcentajes y los puntajes en 3D de las variables más relevantes, a la izquierda los 4 clúster, a la derecha se eliminaron los estados de mayor Trust Media y se observa mejor como el clúster de Menor confianza está más cerca de los ejes coordenados.

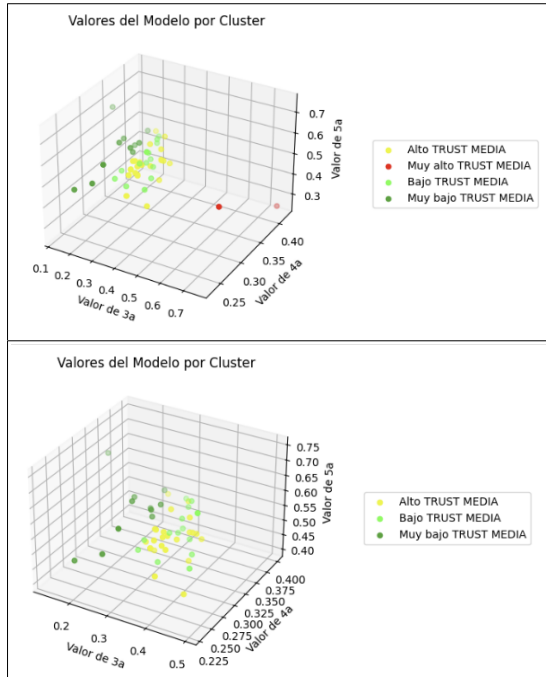


Figura 6. Datos de las variables seleccionadas.

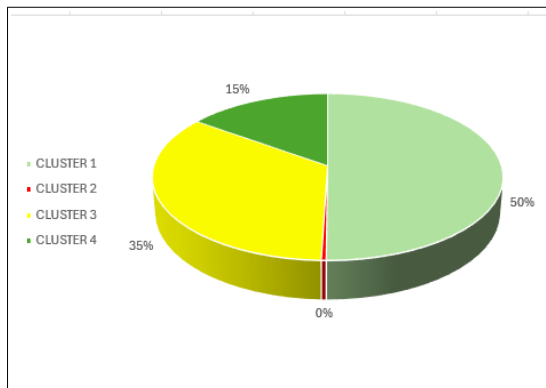


Figura 7. Porcentaje de participantes por Clúster, el clúster 2 tiene un 0.4%.

Contrastando con los reportados en el proyecto de los top 10 en los más confiados y los menos, Fig. 8, vemos que solo Alaska en la clasificación realizada en este estudio salió en el clúster de los No confiados y en lo

reportado en la web del proyecto resultó en 6° lugar de los más confiados.

Alaska pertenece al clúster 1. Como mencionamos anteriormente, en este clúster los estados varían con respecto a la confianza en su capacidad de identificar noticias falsas. En este clúster, Alaska es el estado que mayor porcentaje de participantes que dicen no poder detectar noticias falsas. Por lo tanto, cuenta con las mismas características que el clúster 3 (amarillo).

## 5. DISCUSIÓN Y CONCLUSIONES

### 5.1. Desinformación como patología del sistema cognitivo extendido

Desde la perspectiva de la mente extendida de Andy Clark [42,43], los resultados obtenidos evidencian una ruptura en la integración funcional entre cognición interna y herramientas digitales. Los adolescentes operan en un sistema cognitivo extendido donde redes sociales y motores de búsqueda funcionan como componentes externos de sus procesos de decisión. La clasificación de perfiles de confianza (clusters 1-4) revela distintos grados de patología en este sistema extendido. El cluster 2 (Hawaii, Rhode Island) muestra alta confianza tanto en medios como en capacidad de detección, sugiriendo un sistema extendido funcional pero vulnerable a información corrupta. La desinformación actúa como un 'parásito' que corrompe los elementos externos del sistema cognitivo extendido. El cluster 3 exhibe el patrón opuesto: desconfianza tanto en medios como en su propia capacidad de detección. Esta actitud, aunque los protege de aceptar información falsa, genera una disfunción diferente. Al desconfiar sistemáticamente de todas las fuentes externas, estos adolescentes prescinden de herramientas cognitivas potencialmente útiles. En términos de Clark, esto equivale a amputar componentes externos del sistema cognitivo, reduciendo su capacidad de actuar efectivamente en el mundo.

MOST TRUST MEDIA												
State	1a	1b	2a	2b	3a	3b	4a	4b	4c	5a	5b	Cluster Modelo UNO
Hawaii	0.87	0.13	0.68	0.32	0.75	0.25	0.42	0.5	0.08	0.25	0.75	Cluster Modelo UNO
Rhode Island	0.83	0.17	0.7	0.3	0.61	0.39	0.35	0.57	0.08	0.33	0.67	
Delaware	0.86	0.14	0.8	0.2	0.5	0.5	0.24	0.67	0.09	0.75	0.25	
Wisconsin	0.85	0.15	0.64	0.36	0.47	0.53	0.25	0.69	0.06	0.55	0.45	
Iowa	0.85	0.15	0.57	0.43	0.44	0.56	0.28	0.65	0.08	0.54	0.46	
Alaska	0.8	0.2	0.6	0.4	0.42	0.58	0.28	0.64	0.08	0.4	0.6	
Ohio	0.84	0.16	0.64	0.36	0.41	0.59	0.3	0.63	0.07	0.55	0.45	
Arkansas	0.79	0.21	0.67	0.33	0.39	0.61	0.32	0.59	0.08	0.64	0.36	
South Dakota	0.83	0.17	0.68	0.32	0.39	0.61	0.25	0.66	0.08	0.54	0.46	
Kentucky	0.87	0.13	0.6	0.4	0.39	0.61	0.3	0.61	0.09	0.61	0.39	
LEAST TRUST MEDIA												
State	1a	1b	2a	2b	3a	3b	4a	4b	4c	5a	5b	Cluster Modelo UNO
Vermont	0.86	0.14	0.57	0.43	0.13	0.87	0.35	0.61	0.04	0.67	0.33	Cluster Modelo UNO
North Dakota	0.84	0.16	0.6	0.4	0.2	0.8	0.23	0.69	0.09	0.5	0.5	
Maryland	0.85	0.15	0.68	0.32	0.23	0.78	0.4	0.52	0.08	0.5	0.5	
Washington	0.82	0.18	0.64	0.36	0.24	0.76	0.32	0.62	0.06	0.6	0.4	
Utah	0.77	0.23	0.65	0.35	0.26	0.74	0.3	0.65	0.05	0.62	0.38	
West Virginia	0.84	0.16	0.65	0.35	0.26	0.74	0.24	0.68	0.08	0.53	0.47	
Florida	0.84	0.16	0.72	0.28	0.27	0.73	0.36	0.58	0.07	0.58	0.42	
Arizona	0.79	0.21	0.64	0.36	0.27	0.73	0.32	0.62	0.06	0.46	0.54	
New Hampshire	0.85	0.15	0.64	0.36	0.27	0.73	0.33	0.58	0.09	0.56	0.44	
Connecticut	0.84	0.16	0.68	0.32	0.28	0.72	0.33	0.6	0.07	0.42	0.58	

Figura 8. Semaforización con los top 10 presentados en el proyecto.

La desconfianza absoluta impide que información veraz fluya desde el entorno digital hacia los procesos internos de toma de decisiones. El resultado es apatía informativa: si todas las fuentes son sospechosas, el individuo carece de base para formar intenciones y actuar. Esta ruptura autoimpuesta de la cadena causal entre entorno digital y cognición interna representa otra forma de patología del sistema extendido, donde el problema no es información corrupta sino ausencia de integración funcional entre componentes internos y externos.

Los datos muestran que aproximadamente 50% de adolescentes no confían en su capacidad de identificar información falsa, indicando pérdida de equilibrio cognitivo por falta de fiabilidad y accesibilidad constante de información veraz en su entorno digital. Esta ruptura de la causalidad funcional implica que la intencionalidad se ve comprometida: si la intención de estar informado se basa en fuentes sesgadas o falsas, las acciones subsiguientes estarán

mal dirigidas, socavando el propósito cognitivo de la búsqueda de información. La alfabetización mediática emerge entonces no solo como habilidad educativa, sino como restauración de la integridad del sistema cognitivo extendido, permitiendo que los componentes externos recuperen su función de representar el mundo verazmente para lograr objetivos intencionales.

### 5.2 Conclusiones generales

El objetivo central de este estudio fue aplicar técnicas de *machine learning* (ML) para profundizar en el análisis de los resultados de una encuesta que clasifica el nivel de confianza de los participantes en los medios de comunicación. La confianza en los medios es un elemento clave para el funcionamiento de una democracia saludable, ya que influye en la manera en que los ciudadanos acceden a la información, forman opiniones y participan en el discurso público [26]. No obstante, el aumento de la desconfianza hacia los medios puede

favorecer la desinformación y generar apatía cívica.

Los resultados obtenidos mediante la aplicación de técnicas de ML evidencian que una adecuada selección y ajuste de los parámetros de aprendizaje permite complementar y enriquecer los análisis tradicionales, además de demostrar que los datos pueden generar modelos confiables. En particular, los modelos basados en *Decision Trees* (DT) y *Random Forest* (RF) alcanzaron niveles de precisión elevados. Estos resultados sugieren que las técnicas de aprendizaje supervisado son efectivas para identificar las preguntas más relevantes y para clasificar las respuestas de los participantes. Asimismo, el análisis de clústeres permitió identificar cuatro grupos que representan distintos perfiles de confianza, tanto en las noticias como en la capacidad para detectar información falsa.

La interpretación de estos hallazgos indica que aproximadamente el 15 % de los participantes (clúster 4) no confía plenamente en las noticias, mientras que cerca del 50 % (clúster 1) manifiesta una baja confianza en su capacidad para identificar noticias falsas. Estos resultados ponen de manifiesto que la habilidad de los adolescentes para reconocer información engañosa sigue siendo un tema de creciente interés, especialmente en un entorno digital caracterizado por la abundancia de contenidos y la propagación constante de desinformación.

En este contexto, los hallazgos del estudio respaldan la importancia de fortalecer la alfabetización mediática y el desarrollo de habilidades críticas en los jóvenes. Diversas estrategias orientadas a reducir la desinformación destacan la necesidad de enseñar a los usuarios a evaluar de manera crítica la información que consumen y a reconocer fuentes confiables [27-30]. La educación en alfabetización mediática se presenta, por tanto, como una herramienta

fundamental para fomentar una ciudadanía informada y participativa.

En cuanto a líneas futuras de investigación, se propone ampliar el alcance de la muestra, particularmente en la población hispana —incluyendo el periodo post-COVID—, explorar otros algoritmos de ML y evaluar el impacto de intervenciones educativas específicas en la alfabetización mediática de los jóvenes. Continuar investigando en esta área permitirá generar evidencia empírica que apoye el diseño de políticas públicas y programas educativos orientados a combatir la desinformación.

Respecto a la implementación del ML, el uso de Python y de sus paquetes especializados ofrece múltiples ventajas para el desarrollo y la experimentación con algoritmos de aprendizaje automático. Sin embargo, es fundamental considerar sus limitaciones para garantizar una aplicación efectiva y eficiente en distintos contextos. El empleo de analogías y explicaciones accesibles sobre la cognición artificial puede facilitar la comprensión de estos modelos por parte de personas con menor experiencia técnica, permitiéndoles entender cómo los algoritmos toman decisiones basadas en datos [31, 32]. Además, esta aproximación resalta la flexibilidad de los modelos y la diversidad de resultados posibles, aspectos clave para comprender tanto sus aplicaciones prácticas como sus limitaciones en escenarios reales [33-35].

En conclusión, este estudio pone de relieve la eficacia de las técnicas de ML para el análisis de la confianza en los medios y la detección de noticias falsas entre adolescentes. Los resultados aportan información valiosa para abordar este desafío en la era digital y subrayan la necesidad de continuar investigando y desarrollando estrategias educativas y tecnológicas que fortalezcan la capacidad crítica de los jóvenes frente a la desinformación.

## REFERENCIAS

- [1] Newman N, Fletcher R, Robertson CT, Arguedas AR, Nielsen RK. Reuters Institute digital news report 2024. Oxford: Reuters Institute for the Study of Journalism; 2024. Disponible en: <https://tinyurl.com/yvwbnpyx>. doi: [20.500.12592/kprrc2m](https://doi.org/10.500.12592/kprrc2m)
- [2] Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science*. 2018 Mar 9;359(6380):1146-1151. doi: [10.1126/science.aap9559](https://doi.org/10.1126/science.aap9559)
- [3] Peng RD, Hicks SC. Reproducible Research: A Retrospective. *Annual Review of Public Health*. 2021 Apr 2;42(1):79-93. doi: [10.1146/annurev-publhealth-012420-105110](https://doi.org/10.1146/annurev-publhealth-012420-105110)
- [4] Schackart KE III, Imker HJ, Cook CE. Detailed Implementation of a Reproducible Machine Learning-Enabled Workflow. *Data Science Journal*. 2024;23:1-14. doi: [10.5334/dsj-2024-023](https://doi.org/10.5334/dsj-2024-023)
- [5] Bao W, Lianju N, Yue K. Integration of unsupervised and supervised machine learning algorithms for credit risk assessment. *Expert Systems with Applications*. 2019 Aug 15;128:301-315. doi: [10.1016/j.eswa.2019.02.033](https://doi.org/10.1016/j.eswa.2019.02.033)
- [6] Alpaydin E. *Machine learning*. Cambridge, MA: MIT Press; 2021. doi: [10.7551/mitpress/13811.001.0001](https://doi.org/10.7551/mitpress/13811.001.0001)
- [7] Mitchell TM. Does machine learning really work? *AI Magazine*. 1997 Fall;18(3):11-11. doi: [10.1609/aimag.v18i3.1303](https://doi.org/10.1609/aimag.v18i3.1303)
- [8] Proyecto "Teen Fake News Poll on After School" [Internet]. 2017 [citado 3 ago 2024]. Disponible en: <https://data.world/afterschool/teen-fake-news-poll-on-after-school>
- [9] Persson I, Khojasteh J. Python packages for exploratory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*. 2021;28(6):983-988. doi: [10.1080/10705511.2021.1910037](https://doi.org/10.1080/10705511.2021.1910037)
- [10] Au CH, Ho KK, Chiu DK. The role of online misinformation and fake news in ideological polarization: barriers, catalysts, and implications. *Information Systems Frontiers*. 2022;1-24. doi: [10.1007/s10796-021-10133-9](https://doi.org/10.1007/s10796-021-10133-9)
- [11] Thompson RC, Joseph S, Adeliyi TT. A systematic literature review and meta-analysis of studies on online fake news detection. *Information*. 2022 Nov 1;13(11):527. doi: [10.3390/info13110527](https://doi.org/10.3390/info13110527)
- [12] Pennycook G, Rand DG. The psychology of fake news. *Trends in Cognitive Sciences*. 2021 May;25(5):388-402. doi: [10.1016/j.tics.2021.02.007](https://doi.org/10.1016/j.tics.2021.02.007)
- [13] Bryanov K, Vziatyshva V. Determinants of individuals' belief in fake news: A scoping review determinants of belief in fake news. *PLoS ONE*. 2021 Jun 24;16(6):e0253717. doi: [10.1371/journal.pone.0253717](https://doi.org/10.1371/journal.pone.0253717)
- [14] Herrero-Díaz P, Conde-Jiménez J, Reyes de Cózar S. Teens' motivations to spread fake news on WhatsApp. *Social Media + Society*. 2020 Jul 1;6(3):2056305120942879. doi: [10.1177/2056305120942879](https://doi.org/10.1177/2056305120942879)
- [15] Eraña-Díaz ML, Rosales-Lagarde A. Algoritmos inteligentes aplicados al análisis del perfil de estrés con perspectiva de género. *Journal of Behavior, Health & Social Issues*. 2024;16(1).
- [16] Smith ST, Kao EK, Mackin ED, Shah DC, Simek O, Rubin DB. Automatic detection of influential actors in disinformation networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2021 Jan 26; 118(4): e2011216118. doi: [10.1073/pnas.2011216118](https://doi.org/10.1073/pnas.2011216118)
- [17] Langley D, Reidy C, Towey M, Manisha, Dennehy D. Developing machine learning model for predicting social media induced fake news. In: *Responsible AI and Analytics for an Ethical and Inclusive Digitized Society: 20th IFIP WG 6.11 Conf. on e-Business, e-Services and e-Society, I3E 2021, Galway, Ireland, September 1-3, 2021, Proc. 20*. Springer; 2021. p. 656-669. doi: [10.1007/978-3-030-85447-8\\_55](https://doi.org/10.1007/978-3-030-85447-8_55)
- [18] Varma R, Verma Y, Vijayvargiya P, Churi PP. A systematic survey on deep learning and machine learning approaches of fake news detection in the pre-and post-COVID-19 pandemic. *International Journal of Intelligent Computing and Cybernetics*. 2021;14(4):617-646. doi: [10.1108/IJICC-06-2021-0114](https://doi.org/10.1108/IJICC-06-2021-0114)
- [19] After School [Internet]. [citado 3 ago 2024]. Disponible en: [https://play.google.com/store/apps/details?id=co.april2019.as&hl=es\\_MX&pli=1](https://play.google.com/store/apps/details?id=co.april2019.as&hl=es_MX&pli=1)
- [20] Data World [Internet]. 2019 [citado 3 ago 2024]. Disponible en: <https://www.dataworld.com/es/compania/historia>
- [21] Alikhanian H, Crawford D, DeSouza J, Blohm G. Adaptive cluster analysis approach for functional localization using magnetoencephalography. *Frontiers in Neuroscience*. 2013 May;7:73. doi: [10.3389/fnins.2013.00073](https://doi.org/10.3389/fnins.2013.00073)
- [22] Elhai JD, Montag C. The compatibility of theoretical frameworks with machine learning analyses in psychological research. *Current Opinion in Psychology*. 2020 Aug; 36:1-5. doi: [10.1016/j.copsyc.2020.05.002](https://doi.org/10.1016/j.copsyc.2020.05.002)
- [23] Fuhrer C, Solem JE, Verdier O. *Scientific Computing with Python: High-performance scientific computing with NumPy, SciPy, and pandas*. Birmingham: Packt Publishing Ltd; 2021.
- [24] Ahmed H, Lofstead J. Managing randomness to enable reproducible machine learning. In: *Proceedings of the 5th International Workshop on Practical Reproducible Evaluation of Computer Systems*; 2022 Jun. p. 15-20. doi: [10.1145/3526062.3536354](https://doi.org/10.1145/3526062.3536354)
- [25] Umargono E, Suseno JE, Gunawan SV. K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula. In: *The 2nd International Seminar on Science and Technology (ISSTEC 2019)*. Atlantis Press; 2020. p. 121-129. doi: [10.2991/assehr.k.201010.019](https://doi.org/10.2991/assehr.k.201010.019)
- [26] Hanz K, Kingsland ES. Fake or for real? A fake news workshop. *Reference Services Review*. 2020;48(1):91-112. doi: [10.1108/RSR-09-2019-0062](https://doi.org/10.1108/RSR-09-2019-0062)
- [27] McGrew S, Ortega T, Breakstone J, Wineburg S. The challenge that's bigger than fake news: Teaching students to engage in civic online reasoning.

- American Educator. 2017 Fall;41(3):4. Disponible en: <https://www.common-sense.org/sites/default/files/pdf/2018-05/document-bigger-than-fake-news.pdf>
- [28] McGrew S, Breakstone J, Ortega T, Smith M, Wineburg S. How students evaluate digital news sources. En: Unpacking Fake News: An Educator's Guide to Navigating the Media with Students. 2019. p. 60-73. Disponible en: [https://www.academia.edu/65352037/Unpacking\\_Fake\\_News\\_an\\_Educator\\_s\\_Guide\\_to\\_Navigating\\_the\\_Media\\_with\\_Students](https://www.academia.edu/65352037/Unpacking_Fake_News_an_Educator_s_Guide_to_Navigating_the_Media_with_Students)
- [29] Pangrazio L. What's new about 'fake news'? Critical digital literacies in an era of fake news, post-truth and clickbait. Páginas de Educación. 2018;11(1):6-22. doi: [10.22235/pe.v11i1.1551](https://doi.org/10.22235/pe.v11i1.1551)
- [30] Stein-Smith K. Librarians, information literacy, and fake news. Strategic Library. 2017;37(1-4):1-23. Disponible en: [https://www.libraryspot.net/SL/SL\\_Mar17\\_1.pdf](https://www.libraryspot.net/SL/SL_Mar17_1.pdf)
- [31] Lavelli A, Califf ME, Ciravegna F, Freitag D, Giuliano C, Kushmerick N, Ireson N. Evaluation of machine learning-based information extraction algorithms: criticisms and recommendations. Language Resources and Evaluation. 2008 Dec;42:361-393. doi: [10.1007/s10579-008-9079-3](https://doi.org/10.1007/s10579-008-9079-3)
- [32] Dobson JE. Can an algorithm be disturbed? Machine learning, intrinsic criticism, and the digital humanities. College Literature. 2015 Fall;42(4):543-564. doi: [10.1353/lit.2015.0037](https://doi.org/10.1353/lit.2015.0037)
- [33] Blanco-Justicia A, Sánchez D, Domingo-Ferrer J, Muralidhar K. A critical review on the use (and misuse) of differential privacy in machine learning. ACM Computing Surveys. 2022 Aug;55(8):1-16. doi: [10.1145/3547139](https://doi.org/10.1145/3547139)
- [34] Rawson A, Brito M. A survey of the opportunities and challenges of supervised machine learning in maritime risk analysis. Transport Reviews. 2023;43(1):108-130. doi: [10.1080/01441647.2022.2036864](https://doi.org/10.1080/01441647.2022.2036864)
- [35] Zhong X, Gallagher B, Liu S, et al. Explainable machine learning in materials science. npj Computational Materials. 2022 Sep 23;8:204. doi: [10.1038/s41524-022-00884-7](https://doi.org/10.1038/s41524-022-00884-7)
- [36] Wu Q, Kumar V, Ross QJ, Ghosh J, Motoda H, McLachlan GJ, Ng A, Liu B, Philip SY. Top 10 algorithms in data mining. Knowledge and Information Systems. 2008 Jan;14(1):1-37. doi: [10.1007/s10115-007-0114-2](https://doi.org/10.1007/s10115-007-0114-2)
- [37] Bholowalia P, Kumar A. EBK-means: A clustering technique based on elbow method and k-means in WSN. International Journal of Computer Applications. 2014;105(9).
- [38] Bottou L, Bengio Y. Convergence properties of the k-means algorithms. Advances in Neural Information Processing Systems. 1994;7.
- [39] Pelleg D, Moore AW. X-means: Extending k-means with efficient estimation of the number of clusters. ICML. 2000;1:727-734.
- [40] Chi S, Li X, Tian Y, Li J, Kong X, Ding K, et al. Semi-supervised learning to improve generalizability of risk prediction models. Journal of Biomedical Informatics. 2019 Apr;92:103117. doi: [10.1016/j.jbi.2019.103117](https://doi.org/10.1016/j.jbi.2019.103117)
- [41] Shaik AB, Srinivasan S. A Brief Survey on Random Forest Ensembles in Classification Model. En: Bhattacharyya S, et al., editores. Int. Conf. on Innovative Computing and Communications. LNNS. Singapore: Springer; 2019. p. 273-283. doi: [10.1007/978-981-13-2354-6\\_27](https://doi.org/10.1007/978-981-13-2354-6_27)
- [42] Clark A, Chalmers D. The extended mind. Analysis. 1998 Jan;58(1):7-19. doi: [10.1093/analys/58.1.7](https://doi.org/10.1093/analys/58.1.7)
- [43] Chalmers D. Extended cognition and extended consciousness. Andy Clark and his critics. 2019;9-20. <https://philarchive.org/archive/CHAECA-9>

## ACERCA DE LOS AUTORES



La Dra. Marta Lilia Eraña Díaz es Doctora en Ingeniería y Ciencias Aplicadas por la UAEMor. Cuenta con una licenciatura en Matemáticas Aplicadas a la Computación (UAM-I) y dos maestrías: en Ingeniería Industrial (ITESM CCM) y en Ciencias Cognitivas (CITPSI-UAEMor). Actualmente es profesora investigadora en el Centro de Investigación en Ciencias Cognitivas (CINCCO), donde también imparte clases en la Facultad de Ciencias Químicas e Ingeniería. Es miembro del Sistema Nacional de Investigadores (SNII) y miembro senior del IEEE, participando en la Computer Society y el grupo de afinidad WIE. Su trayectoria incluye liderar proyectos de automatización tecnológica alineados a la Industria 4.0. Colabora en el desarrollo de modelos matemáticos para análisis de datos en el Instituto de Psiquiatría Ramón de la Fuente Muñiz y el Instituto Profesional Multidisciplinario S.C., integrando disciplinas como ingeniería, cognición y análisis de datos.



El Dr. Jorge Oseguera Gamba realizó su licenciatura en filosofía en la Universidad Nacional Autónoma de México, donde también obtuvo un diplomado en

bioética. Llevó a cabo un intercambio en University of California, Berkeley. Recibió una beca Fulbright-García Robles para realizar su maestría y su doctorado en Florida State University, donde ha dictado clases, así como en University of Colorado, Boulder. Actualmente es profesor investigador en el Centro de Investigación en Ciencias Cognitivas (CINCCO) de la Universidad Autónoma del Estado de Morelos (UAEM), donde también imparte clases sobre Políticas Públicas. Es miembro del SNI. Sus principales líneas de investigación son bienestar y cognición política. Actualmente sus proyectos en estas áreas son: 1) desarrollar un marco teórico y conceptual para investigar el bienestar de manera transdisciplinar y 2) analizar los impactos que la Inteligencia Artificial tendrá en la democracia.



La Dra. Nadia Lara Ruiz, realizó su Doctorado en Ingeniería y Ciencias Aplicadas con especialización en Tecnología de Materiales por la UAEMor.

Su investigación doctoral se centró en desarrollar un prototipo de cámara para acelerar el crecimiento de plántulas, aportando soluciones agrícolas frente al cambio climático. Con una Maestría y Licenciatura en Ingeniería Industrial (FCQeI-UAEMor), ha contribuido a investigaciones, abordando mejoras en procesos agrícolas y su impacto en los Objetivos de Desarrollo Sostenible. Además, ha impartido cursos innovadores sobre el uso de inteligencia artificial en la docencia, colaborando con diversas instituciones educativas. Recientemente ha centrado su investigación en el impacto de la inteligencia artificial en la docencia, destacándose por su compromiso con la enseñanza, la investigación y la innovación tecnológica en agricultura y educación.

#### APÉNDICE A.

Proyecto "Teen Fake News Poll on After School"

<https://data.world/afterschool/teen-fake-news-poll-on-after-school>

#### DESCRIPCIÓN

Se preguntó a los adolescentes estadounidenses qué saben y qué piensan sobre las "noticias falsas". After School, la red social más grande centrada en adolescentes encuestó a sus usuarios sobre el tema de las noticias falsas. Durante varios días, decenas de miles de estudiantes de secundaria en los 50 estados participaron en la encuesta. De los más de 39,000 estudiantes que respondieron a la pregunta "¿has oído hablar de noticias falsas?" El 21% de los adolescentes nunca había oído hablar de esto.

En la publicación realizada por Katie Notopoulos en BuzzFeed News, comenta:

"Queríamos encuestar a algunos estudiantes reales de secundaria, y la aplicación social After School verifica que sus usuarios estén realmente en la escuela secundaria a través de su Facebook y otros factores. Los adolescentes de los 50 estados respondieron las preguntas de la encuesta realizadas en la aplicación: poco más de 39 mil adolescentes en total.

Un estudio de Stanford del año 2016 mostró que los estudiantes de secundaria y preparatoria no son muy buenos para determinar noticias falsas, especialmente cosas más matizadas como notar sesgos en una fuente o comprender la diferencia entre contenido patrocinado y un artículo normal"

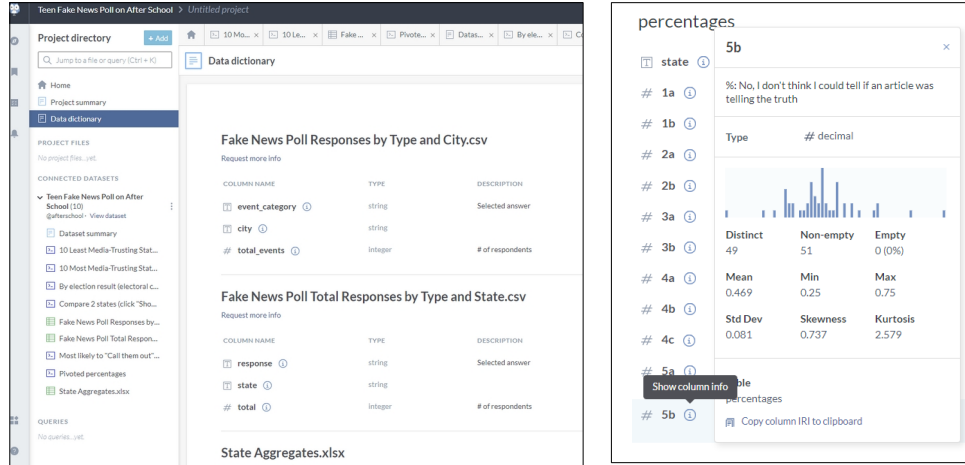


Figura 8. Imágenes de la plataforma del proyecto.

Se publicaron tres bases de datos:

ID	Nombre base de datos	Tipo	Descripción	Columnas	Renglones
BD1	Fake News Poll Responses by Type and City.csv	csv	Respuestas a las 5 preguntas, (11 en total) por ciudades	3	34,321
BD2	Fake News Poll Total Responses by Type and State.csv	csv	Respuestas a las 5 preguntas (11 en total) por estado	3	572
BD3	State Aggregates.xlsx	xlsx	Suma de respuestas por pregunta por estado para cada respuesta en hoja diferente, la sexta hoja de cálculo tiene los porcentajes de respuestas de cada pregunta	Hoja 1 a la 5 con 3 columnas, Hoja 6 con 11 columnas	52

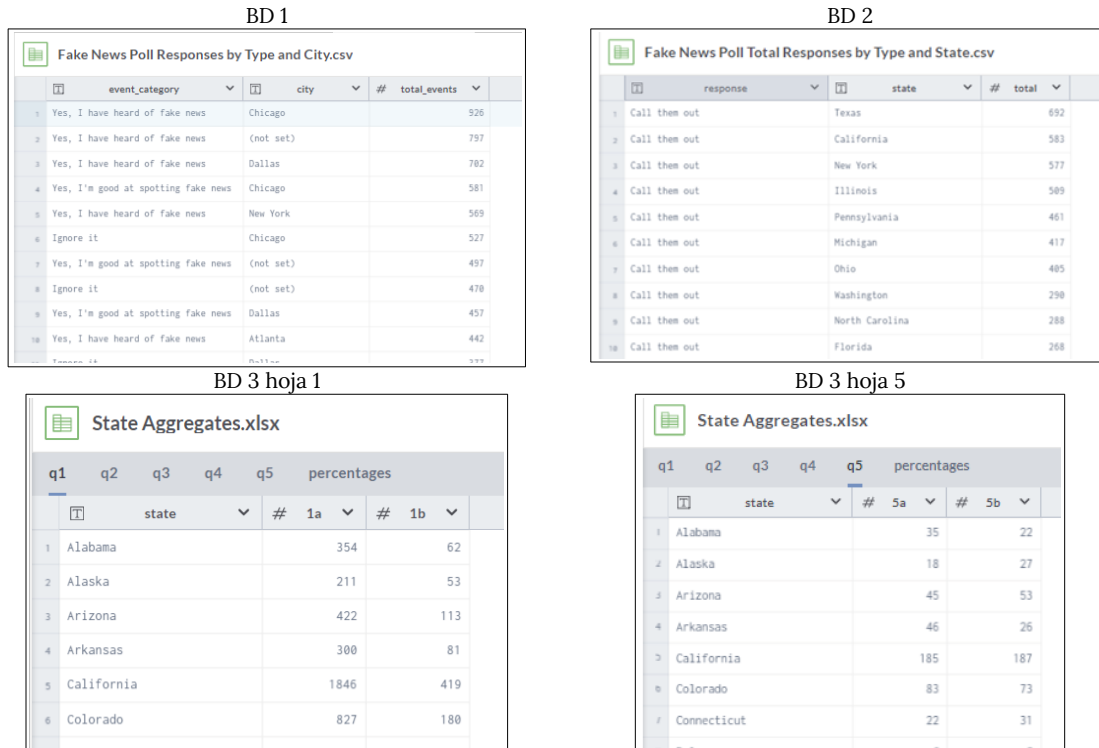


Figura 9. Imagen de las bases de datos.

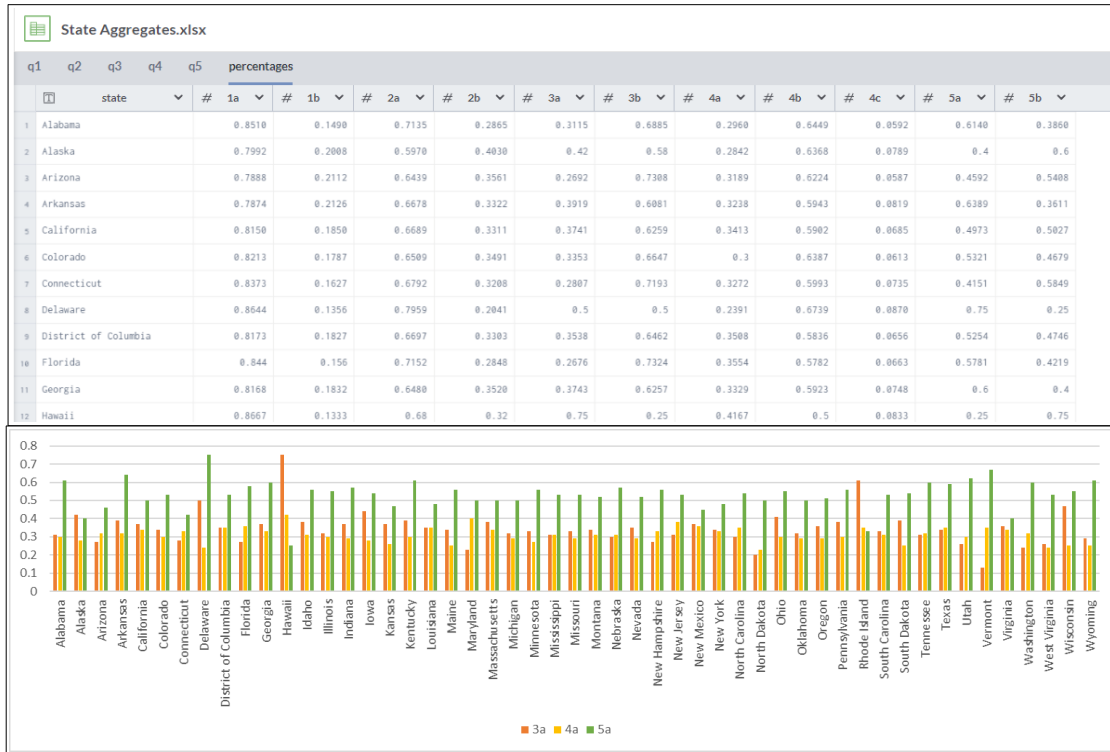


Figura 10. Resultados por estado a las preguntas relevantes del estudio con ML.

En la figura 10 se pueden ver los porcentajes de las respuestas más relevantes para el análisis de ML, se observa que los estados de Hawaii y Rhode Island tienen mayor puntuación en la respuesta 3a y estos están clasificados en el clúster de mayor TM.

### APÉNDICE B.

Algoritmos utilizados en el estudio.

Modelos del aprendizaje máquina, ML, para agrupación y clasificación

Los principios de funcionamiento de cada algoritmo de aprendizaje automático se describen en las siguientes subsecciones.

**K-Means:** Se ha identificado como una de las clasificaciones de aprendizaje no supervisado más populares [36].

Es un algoritmo de aprendizaje no supervisado que particiona el conjunto de objetos en un determinado número de grupos de acuerdo con un criterio de optimización. En su implementación la convergencia a una agrupación óptima no está garantizada y para problemas grandes, se requiere de un mayor número de iteraciones. Para determinar el número de grupos o clústeres, se utiliza el algoritmo de Elbow [37].

De acuerdo con la literatura especializada [38], [39] se pueden identificar cuatro pasos en el algoritmo:

- Paso 1. Inicialización: Se definen un conjunto de objetos a particionar, el número de grupos y un centroide por cada grupo. Algunas implementaciones del algoritmo estándar determinan los centroides iniciales de forma aleatoria; mientras que algunos otros procesan los datos y determinan los centroides mediante cálculos.
- Paso 2. Clasificación: Para cada objeto de la base de datos, se calcula su distancia a cada centroide, se determina el centroide más cercano, y el objeto es incorporado al grupo relacionado con ese centroide.
- Paso 3. Cálculo de centroides: Para cada grupo generado en el paso anterior se vuelve a calcular su centroide.
- Paso 4. Condición de convergencia: Se han usado varias condiciones de convergencia, de las cuales las más utilizadas son las siguientes: converger cuando alcanza un número de iteraciones dado, converger cuando

no existe un intercambio de objetos entre los grupos, o converger cuando la diferencia entre los centroides de dos iteraciones consecutivas es más pequeña que un umbral dado. Si la condición de convergencia no se satisface, se repiten los pasos dos, tres y cuatro del algoritmo.

Supongamos que el número de grupos sea  $K=3$ , la formulación matemática es un conjunto de  $n$  datos o instancias en el vector  $d$ -dimensional  $D=\{x_1, \dots, x_n\} \in \mathbb{R}^d$ , donde el problema es determinar un conjunto de 3 centroides  $C=\{c_1, c_2, c_3\} \in \mathbb{R}^d$  donde se minimice la función del error, ecuación A1).

$$E(C) = \sum_{x \in D} \|x - C_k\|^2 \quad A1)$$

Consiste en encontrar la partición de las  $n$  instancias en 3 grupos cuya distancia al centro sea la mínima.

Una métrica para evaluar el agrupamiento es el ratio que se obtiene al dividir la suma de cuadrados del error dentro del grupo (WCSS, Within-Cluster Sum of Squares) entre la suma de cuadrados total (TSS, Total Sum of Squares). Se conoce comúnmente como la "proporción de varianza explicada" o "ratio de varianza explicada". Si el ratio es cercano a 1: Esto indica que los clusters son efectivos para capturar la estructura de los datos, y la varianza en los datos se explica principalmente por la agrupación en clusters. En este caso, los clusters son representativos de las similitudes intrínsecas en los datos. Si el ratio es cercano a 0: Esto sugiere que los clusters no están capturando bien la estructura de los datos y que la mayoría de la varianza en los datos no se explica mediante la agrupación en clusters. Puede ser un indicio de que el número de clusters ( $k$ ) es demasiado bajo o que el algoritmo no ha logrado encontrar una buena agrupación.

**Árbol de decisión, DT:** El aprendizaje supervisado de árboles de decisión es una de las técnicas de inferencia inductiva más usadas [40], [5]. Se trata de un método para aproximar funciones de valores discretos que sea capaz de expresar hipótesis disyuntivas y que presenten capacidad en la discriminación del ruido de los datos en los ejemplos de entrenamiento.

De manera simple, los árboles de decisión permiten representar en forma gráfica una serie de reglas sobre la decisión que debe tomarse en la asignación de un valor de salida a un registro dado, se usan comúnmente en el análisis para toma de decisión, ayudando a identificar los elementos relevantes para alcanzar una meta. Al crear un árbol completo que ha clasificado todo el conjunto de datos, una medida de "bondad", es decir, qué tan bien los atributos se discriminan entre las clases, se empleó el índice de Gini, esta función mide la impureza "de un atributo con respecto a las clases". Dadas las probabilidades para cada clase  $p_i$ , la función general de Gini, o medida de impureza se presenta en la ecuación A2 donde la impureza de Gini se calcula sumando los productos por pares de estas probabilidades para cada etiqueta.

$$\sum (p_i \sum p_j) = \sum p_i(1 - p_i) = \sum (p_i - p_i^2) = 1 - \sum p_i^2 \quad A2)$$

**Random Forest o Bosque aleatorio, RF:** El clasificador de RF crea múltiples árboles de decisión a partir de un subconjunto seleccionado aleatoriamente del conjunto de datos de entrenamiento, posteriormente promedia los resultados de diferentes árboles de decisión. En teoría, una gran cantidad de árboles no correlacionados crearán predicciones más precisas que un árbol de decisión individual [41].

Dentro de las métricas para determinar un buen desempeño está la exactitud, ecuación A3, que se obtiene de la clasificación correcta entre el total de elementos a clasificar.

$$Accuracy = \frac{\text{suma de correctas}}{\text{suma correctas e incorrectas}} \quad A3)$$