

SISTEMA DE ALERTA TEMPRANA PARA MONITOREAR LA PROPAGACIÓN DE NOTICIAS FALSAS: CASO DE ESTUDIO CONTEXTO POLÍTICO COLOMBIANO



Universidad
del Cauca

Trabajo de Grado

**Laura Sofía Daza Rosero
César Enrique Manzano Velasco**

Director - Universidad del Cauca:
César Alberto Collazos Ordóñez

Co-Director - Pontificia Universidad Católica de Chile:
Marcelo Gabriel Mendoza Rocha

Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Sistemas
Grupo IDIS
Popayán, Enero 2024

*Dedicado a cada miembro de nuestras familias
el mejor impulso para salir adelante
y alcanzar cada meta propuesta.*

CONTENIDO

LISTA DE TABLAS	V
LISTA DE FIGURAS	V
LISTA DE ACRÓNIMOS	X
1. FASE DE FORMULACIÓN DEL PROBLEMA DE INVESTIGACIÓN	1
1.1. Introducción	2
1.2. Planteamiento del problema	3
1.3. Pregunta de investigación	6
1.4. Objetivo General	6
1.5. Objetivos Específicos	7
1.6. Metodología	7
1.6.1. Investigación Preliminar	8
1.6.2. Diseño y Análisis:	9
1.6.3. Desarrollo:	10
1.6.4. Validación:	10
1.7. Aportes de la investigación	11
1.8. Estructura del documento	12
2. FASE DE ANÁLISIS DE TRABAJOS RELACIONADOS	14
2.1. Introducción	15
2.2. Revisión de literatura y trabajos relacionados	15
2.2.1. Marco teórico	15
Conceptualización de noticias falsas:	16
Características de las noticias falsas:	17
Conceptos relacionados con las noticias falsas:	18
Conceptos implementados en la investigación:	19
2.2.2. Trabajos relacionados	22
2.3. Brechas existentes	28
3. FASE DE DISEÑO Y ANÁLISIS	30
3.1. Introducción	31
3.2. Fase de diseño y análisis	32
3.2.1. Planeación de la fase	32
3.2.2. Desarrollo del modelo para detectar la veracidad de noticias	32

4. FASE DE DESARROLLO DEL MODELO CONCEPTUAL	49
4.1. Introducción	50
4.2. Estructura general	50
4.3. Almacenamiento	53
4.4. Registro de petición de tweets	56
4.5. Validación de tweets(Fine tuning)	57
4.6. Notificación a los usuarios:	59
4.7. App de registro y visualización	64
5. FASE DE VALIDACIÓN DEL MODELO CONCEPTUAL	67
5.1. Introducción	68
5.2. Prueba de validación del prototipo	68
5.3. Ejecución de la prueba de validación	69
5.4. Análisis de resultados	71
6. CONCLUSIONES Y TRABAJOS FUTUROS	77
6.1. Introducción	78
6.2. Conclusiones	79
6.3. Limitaciones del trabajo	80
6.4. Trabajos futuros	81
REFERENCIAS	88
A. CONSENTIMIENTO INFORMADO PARA LA REALIZACIÓN DE PRUEBAS	89
B. ENCUESTA CSUQ	94
C. RESULTADOS GRÁFICOS DE LAS ENCUESTAS	96

LISTA DE TABLAS

1.1. <i>Perfil número 1</i>	4
1.2. <i>Perfil número 2</i>	5
1.3. <i>Perfil número 3</i>	5
1.4. Cumplimiento de objetivos en cada una de las etapas	7
3.1. <i>Datos seleccionados para cada cuenta</i>	39
3.2. <i>Datos obtenidos</i>	40
3.3. <i>Campos específicos solicitados</i>	41
3.4. Resumen de Tweets.	41
3.5. Resumen de Valor Esperado y Cantidad de Impresiones por Usuario.	43
3.6. Primeras 5 cuentas seleccionadas.	45
3.7. Distribución de tweets según intervalo de valor.	46
3.8. Análisis de Veracidad de Noticias.	47

LISTA DE FIGURAS

1.1. Mapa de contenido capítulo 1 (Figura propia)	2
1.2. Estructura de la monografía (Figura propia)	13
2.1. Mapa de contenido capítulo 2 (Figura propia)	15
2.2. Estructura del prototipo (Figura tomada de 41)	21
3.1. Mapa de contenido capítulo 3 (Figura propia)	31
3.2. Modelo empleado para el desarrollo de la investigación (Figura propia)	32
3.3. Distribución de entrenamiento del modelo (Figura propia)	35
3.4. Proceso de entrenamiento del modelo (Figura propia)	37
3.5. Diagrama de consulta	40
3.6. Gráfica de distribución normal (Figura propia)	43
3.7. Impacto relativo de tweets (Figura propia)	45
4.1. Mapa de contenido capítulo 4 (Figura propia)	50
4.2. Diagrama de componentes del sistema (Figura propia)	51
4.3. Diagrama de secuencia (Figura propia)	52
4.4. Opciones de almacenamiento (Figura propia)	53
4.5. Diagrama de almacenamiento en MongoDB (Figura propia)	54
4.6. Diagrama de almacenamiento en Firebase (Figura propia)	55
4.7. Proceso de validación de tweets (Figura propia)	57
4.8. Script del proceso de verificación (Figura propia)	58
4.9. Sistema de notificaciones mediante SNS y AWS (Figura propia)	59
4.10. Sistema de notificaciones mediante GMAIL (Figura propia)	60
4.11. Sistema de notificaciones mediante SMS (Figura propia)	60
4.12. Diagrama de clases de servicios integrados en la aplicación con FastAPI.	61
4.13. Etiquetas implementadas en FastAPI (Figura propia)	61
4.14. Endpoint de la etiqueta "users"(Figura propia)	62
4.15. EndPoints de la etiqueta "UserTweets" (Figura propia)	62
4.16. EndPoints de la etiqueta "Tweets" (Figura propia)	63
4.17. EndPoints de la etiqueta "News" (Figura propia)	63
4.18. EndPoints de la etiqueta "Notificaciones" (Figura propia)	64

4.19. Interfaz del home de la app (Figura propia)	64
4.20. Interfaz de interacción de App (Figura propia)	65
4.21. Interfaz de registro de información (Figura propia)	66
4.22. Interfaz de opciones disponibles en la app (Figura propia)	66
5.1. Mapa de contenido capítulo 5 (Figura propia)	68
5.2. Estructura de la prueba de validación	69
5.3. Ejemplo de noticias con etiquetas	70
5.4. Ejemplo de notificaciones enviadas	71
5.5. Gráfico de bigotes Características de aplicativo de visualización (Figura propia)	72
5.6. Gráfico de bigotes Evaluación de características del sistema de alerta temprana (Figura propia)	73
5.7. Gráfica de resultados para la pregunta "¿Cómo evalúa la efectividad del sistema de alerta temprana en comparación con otros métodos de combatir la desinformación?" (Figura propia)	75
6.1. Mapa de contenido capítulo 6 (Figura propia)	79
B.1. Cuestionario CSUQ	95
C.1. Gráfica de resultados para la afirmación "En general, estoy satisfecho con lo fácil que es utilizar este sistema" (Figura propia)	97
C.2. Gráfica de resultados para la afirmación "Fue simple usar este sistema" (Figura propia)	97
C.3. Gráfica de resultados para la afirmación "Me siento cómodo utilizando este sistema" (Figura propia)	98
C.4. Gráfica de resultados para la afirmación "Fue fácil aprender a utilizar este sistema" (Figura propia)	98
C.5. Gráfica de resultados para la afirmación "Creo que entendí rápidamente la finalidad de este sistema" (Figura propia)	99
C.6. Gráfica de resultados para la afirmación "La información que proporciona el sistema fue efectiva ayudándome a completar las tareas" (Figura propia)	99
C.7. Gráfica de resultados para la afirmación "La interfaz del sistema fue placentera" (Figura propia)	100
C.8. Gráfica de resultados para la afirmación "Me gustó utilizar el sistema" (Figura propia)	100
C.9. Gráfica de resultados para la afirmación "El sistema tuvo todas las herramientas que esperaba que tuviera" (Figura propia)	101
C.10. Gráfica de resultados para la afirmación "En general, estuve satisfecho con el sistema" (Figura propia)	101

C.11.Gráfica de resultados para la pregunta "¿Cómo calificaría su capacidad para identificar noticias falsas antes de usar el sistema?"(Figura propia)	102
C.12.Gráfica de resultados para la pregunta "¿Notó alguna mejora en su capacidad para identificar noticias falsas después de usar el sistema de alerta temprana?"(Figura propia)	102
C.13.Gráfica de resultados para la pregunta "¿Influyó el sistema en su motivación para buscar más información sobre las noticias después de recibir las notificaciones, especialmente en términos de verificar la veracidad de las noticias?"(Figura propia)	103
C.14.Gráfica de resultados para la pregunta "¿Hubo alguna noticia que inicialmente creyó que era falsa y luego el sistema identificó como verdadera (Visceversa)?"(Figura propia)	103
C.15.Gráfica de resultados para la pregunta "¿Cómo percibe la rapidez de respuesta de un sistema de alerta temprana en la detección de desinformación?"(Figura propia)	104
C.16.Gráfica de resultados para la pregunta "¿Cree que la personalización de las notificaciones, adaptadas a sus intereses, aumentaría la efectividad del sistema de alerta temprana?"(Figura propia)	104
C.17.Gráfica de resultados para la pregunta "¿Considera que la participación activa de la comunidad de usuarios es esencial para mejorar la eficacia del sistema de alerta temprana?"(Figura propia)	105
C.18.Gráfica de resultados para la pregunta "¿Opina que la transparencia en el funcionamiento del sistema de alerta temprana aumentaría su confianza en sus resultados?"(Figura propia)	105
C.19.Gráfica de resultados para la pregunta "¿Cómo evalúa la efectividad del sistema de alerta temprana en comparación con otros métodos de combatir la desinformación?"(Figura propia)	106
C.20.Gráfica de resultados para la pregunta "¿Hasta qué punto le resulta efectivo un sistema de alerta temprana en la lucha contra la desinformación mediante la implementación de notificaciones?"(Figura propia)	106
C.21.Gráfica de resultados para la pregunta "¿En qué medida confía en la capacidad de un sistema de alerta temprana para detectar y combatir la desinformación?"(Figura propia)	107
C.22.Gráfica de resultados para la pregunta "¿Considera que la rapidez de las notificaciones afecta la eficacia del sistema de alerta temprana en combatir la desinformación?"(Figura propia)	107
C.23.Gráfica de resultados para la pregunta "¿Cree que la implementación de notificaciones puede influir en su percepción y comprensión de la información?"(Figura propia)	108

C.24.Gráfica de resultados para la pregunta "¿Cuál sería la frecuencia ideal de notificaciones para mantener informado al usuario y combatir la desinformación de manera efectiva?"(Figura propia)	108
C.25.Gráfica de resultados para la pregunta "¿Piensa que la efectividad del sistema de alerta temprana dependería de la colaboración entre plataformas, usuarios y autoridades?"(Figura propia)	109

LISTA DE ACRÓNIMOS

API *Application Programming Interface.* 8, 10, 36, 39–41, 51–54, 56

AWS *Amazon Web Services.* vi, 51, 59

CRUD *Create, Read, Update and Delete.* 61, 62

CSUQ *Computer System Usability Questionnaire.* IV, VII, 22, 71, 94, 95

MIT *Massachusetts Institute of Technology.* 4

SMS *Short Message Service.* vi, 51, 53, 59–61, 70

SNS *Simple Notification Service.* vi, 51, 53, 59

CAPÍTULO 1

FASE DE FORMULACIÓN DEL PROBLEMA DE INVESTIGACIÓN



1.1. Introducción

En este capítulo, se presenta el tema de investigación tratado en el trabajo de grado, se expone el planteamiento del problema y los objetivos perseguidos. Adicionalmente, se plantea la metodología empleada en la investigación, haciendo énfasis en su aplicabilidad y sus aportes. Finalmente, se describe la estructura de este documento.

De este modo, la Figura 1.1, presenta el contenido abordado en el Capítulo 1:

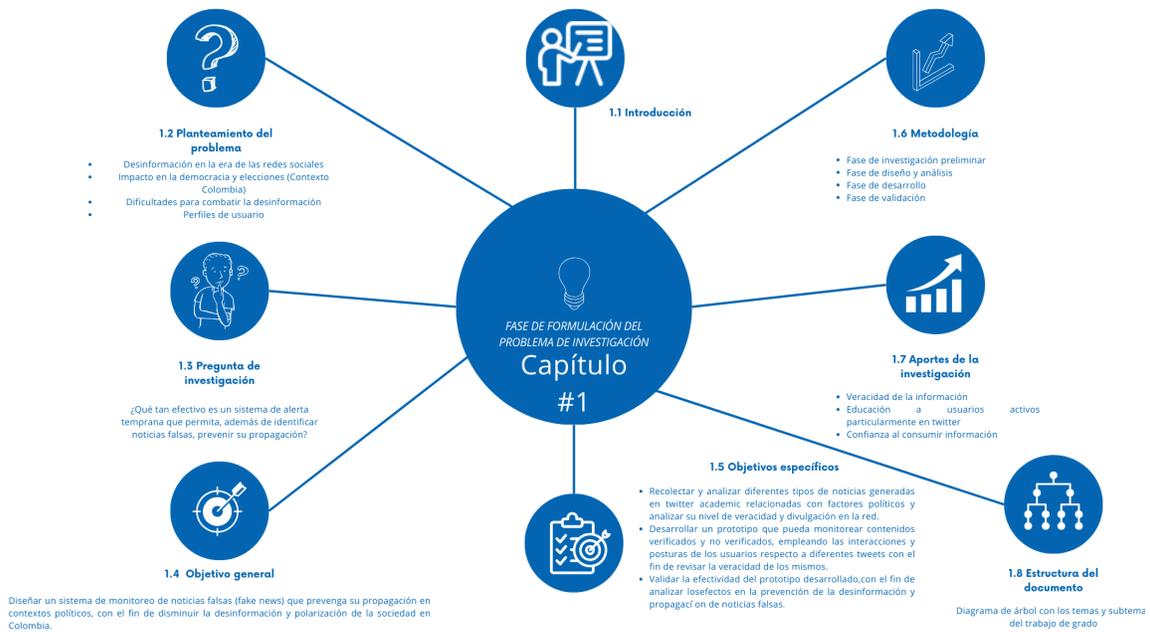


Figura 1.1: Mapa de contenido capítulo 1 (Figura propia)

1.2. Planteamiento del problema

Con el auge de las redes sociales y la facilidad para difundir información a través de ellas, es común encontrar noticias falsas que circulan por internet y se propagan rápidamente, generando desinformación y afectando la percepción de las personas sobre diferentes temas [1].

Es imprescindible enfocar la atención en los efectos perjudiciales de esta proliferación de desinformación, que inciden directamente en la percepción que las personas tienen sobre una amplia gama de asuntos. Dicha desinformación no solo trastoca la ecuanimidad de la esfera informativa, sino que también compromete la capacidad de discernimiento de los individuos al influir en sus decisiones y juicios [2]. La transformación de las redes sociales en canales relevantes para la obtención de noticias y la consecuente adopción de estas plataformas como fuentes primarias de información, como revela el informe de la Fundación Thomson Reuters [3], denota un cambio sustancial en el panorama mediático y en la forma en que el público se involucra con las noticias.

El informe mencionado resalta la evolución marcada en los hábitos informativos de la sociedad contemporánea. El incremento del 47 % al 51 % [4] en el uso de redes sociales como fuente de noticias en tan solo un año pone en evidencia una tendencia ascendente en la preferencia por estos canales, en detrimento de las fuentes tradicionales. Esta transición hacia la virtualidad informativa no solo implica un cambio en la forma en que las personas se mantienen informadas, sino que también instaaura un contexto propicio para la difusión indiscriminada de contenido falso y engañoso [5].

En particular en el contexto de la investigación, las consecuencias pueden ser graves, ya que afectan las decisiones de las personas, influyen en las elecciones, en la opinión pública, y generan pánico o alarma social, asimismo, una vez que se propagan; son muy difíciles de erradicar, lo que dificulta el trabajo de los medios de comunicación y las autoridades para combatir la desinformación [6].

Las ramificaciones de estas consecuencias revisten un carácter significativamente serio, dado que repercuten de manera sustancial en el proceso de toma de decisiones individuales y colectivas. Además, ejercen una influencia de gran alcance en la configuración de las preferencias electorales y la formación de la percepción pública en diversos asuntos de relevancia [1].

Dicho de otra manera, la naturaleza persistente de las informaciones erróneas agrava el panorama, ya que su propagación tiende a arraigarse con tenacidad en la conciencia pública que conlleva a una considerable complejidad en el empeño por erradicar tales narrativas distorsionadas, planteando un desafío sustancial tanto para los medios de comunicación como para las autoridades. La labor de los medios de comunicación se ve comprometida al tener que desmontar y corregir información incorrecta, mientras que las autoridades se enfrentan a un arduo

trabajo para contrarrestar la desinformación en sus múltiples formas y manifestaciones.

En definitiva, es evidente que el impacto de estas situaciones trasciende con creces la esfera de la simple información inexacta, afectando los fundamentos mismos de una sociedad informada y comprometida. Por ende, la necesidad de desarrollar estrategias eficaces para mitigar y contrarrestar estas consecuencias se presenta como una tarea impostergable para garantizar la integridad del discurso público y la salud de la esfera democrática.

Ahora bien, en el contexto político Colombiano en épocas de elecciones, es uno de los mayores desafíos, pues la difusión de información tergiversada en medios digitales y otros canales de comunicación influyen en el resultado de las elecciones y manipulan la opinión pública [7].

Por consiguiente, la información falsa puede tener un impacto negativo en los procesos democráticos al erosionar la confianza en las instituciones, los medios de comunicación y los procesos electorales [8]. En Colombia, este problema es particularmente relevante debido a la historia del conflicto armado interno, la violencia política y la polarización social [9].

De acuerdo a un estudio realizado por investigadores del MIT y de la Sloan School of Management [10], en redes sociales como twitter la información falsa es mas divulgada que la verdadera [11], observando que la probabilidad de que un tuit genere una cascada de retuits es más alta para contenido falso, además este contenido puede ser compartido en otras redes sociales y sumado al alto grado de incidencia de otras plataformas con un uso creciente en Colombia [12], genera grandes perjuicios a la sociedad en materia informativa que no se pueden medir de forma adecuada.

Por tanto, como apoyo a la investigación, para identificar, empatizar y entender a los actores principales dentro de la problemática, particularmente en la red social Twitter, se realizó una encuesta; donde, a partir de un análisis de resultados sobre el comportamiento común de los actores, se construyeron los siguientes perfiles:

Nota: Para conservar la rigurosidad de la encuesta, se han asignado nombres ficticios para los perfiles creados

Tabla 1.1: Perfil número 1

Usuario	Información asociada al perfil
Nombre y edad	Pedro Pérez de 23 años de edad.

Usuario	Información asociada al perfil
Descripción	Usuario activo en la red social Twitter, que dedica entre 15 y 20 minutos diarios para navegar en la red social desde su dispositivo celular. Sin embargo, el tiempo que dedica a leer la información, noticias o contenido, es inferior a 5 minutos.
Descripción de la actividad	La intención de acceder a Twitter es ver, compartir, reaccionar e informarse sobre noticias y otros contenidos. Sus criterios para divulgar la información son: que sea información de actualidad, que provenga de una fuente confiable, que sea un tema global y coincida con otras fuentes o que tenga muchas reacciones por parte de usuarios reconocidos.

Tabla 1.2: Perfil número 2

Usuario	Información asociada al perfil
Nombre y edad	Ricardo Suárez de 27 años de edad.
Descripción	Usuario activo en la red social Twitter, que dedica entre 20 y 60 minutos diarios para navegar en la red social desde su dispositivo celular.
Descripción de la actividad	La intención de acceder a Twitter es únicamente ver e informarse sobre noticias y otros contenidos. Sin embargo, en algunas ocasiones reacciona a algún tweet solo porque le ha llamado la atención.

Tabla 1.3: Perfil número 3

Usuario	Información asociada al perfil
Nombre y edad	Liliana Suárez de 17 años de edad.

Usuario	Información asociada al perfil
Descripción	Usuario activo en la red social Twitter, que dedica entre 60 y 120 minutos diarios para navegar en la red social desde su dispositivo celular. Pero en caso de encontrar una noticia, el tiempo que dedica a leer la información, es inferior a 5 minutos.
Descripción de la actividad	La intención principal de navegar en Twitter es informarse sobre temas de actualidad, moda, tendencias del mercado y tal vez, según la ocasión leer una que otra noticia. En caso de encontrar una noticia, SÓLO teniendo en cuenta que es un tema de su gusto, la comparte, reacciona y difunde la información sin previa verificación.

Por lo expuesto, surge una problemática alrededor de la existencia de noticias falsas y de mecanismos que permitan identificarlas. Esto implica no solo desarrollar estrategias para detectarlas y desmentirlas, sino también educar a las personas para que sean críticas con la información que reciben y puedan discernir entre lo verdadero y lo falso.

1.3. Pregunta de investigación

En consecuencia, este trabajo se encamina a responder la siguiente pregunta de investigación:

¿Qué tan efectivo es un sistema de alerta temprana que permita, además de identificar noticias falsas, prevenir su propagación?

Nota: La efectividad en el sistema se toma en términos del comportamiento de un usuario cuando se enfrenta a una noticia.

1.4. Objetivo General

Diseñar un sistema de monitoreo de noticias falsas (fake news) que prevenga su propagación en contextos políticos, con el fin de disminuir la desinformación y polarización de la sociedad en Colombia.

1.5. Objetivos Específicos

1. Recolectar y analizar diferentes tipos de noticias generadas en twitter academic relacionadas con factores políticos y analizar su nivel de veracidad y divulgación en la red.
2. Desarrollar un prototipo que pueda monitorear contenidos verificados y no verificados, empleando las interacciones y posturas de los usuarios respecto a diferentes tweets con el fin de revisar la veracidad de los mismos
3. Validar la efectividad del prototipo desarrollado, con el fin de analizar los efectos en la prevención de la desinformación y propagación de noticias falsas.

Efectividad: Capacidad de lograr el efecto que se desea o se espera [13]. En el contexto de la investigación, la efectividad fue evaluada en términos del comportamiento de un usuario cuando se enfrenta a una noticia.

1.6. Metodología

Con el fin de alcanzar los objetivos planteados, se empleó la metodología de Investigación y Desarrollo (I&D), esta se centra en la investigación científica y tecnológica con el propósito de diseñar soluciones innovadoras y efectivas [14]. La metodología de I&D se compone de cuatro etapas, en cada una de las cuales se abordan distintos aspectos y objetivos del proyecto.

El orden de ejecución es el siguiente:

Etapas	Cumplimiento de objetivo específico
Investigación Preliminar	número 1
Diseño y Análisis, Desarrollo	número 2
Validación	número 3

Tabla 1.4: Cumplimiento de objetivos en cada una de las etapas

A continuación, se exponen las actividades correspondientes a cada una de las fases de la metodología de I&D implementada en la investigación.

1.6.1. Investigación Preliminar

Esta fase comprendió diversas actividades, tales como:

- Revisión de las diferentes técnicas y metodologías existentes para el análisis de noticias y verificación de la información en redes sociales.
- Análisis de las fuentes de información relevantes en Twitter relacionadas con temas políticos en Colombia.
- Definición los criterios y características de los usuarios que serán incluidos en el análisis de las noticias.
- Análisis de los posibles riesgos y perjuicios que puede causar la difusión de noticias falsas en la sociedad.
- Definición los métodos y tecnologías para la recolección, almacenamiento y análisis de la información de Twitter, considerando herramientas como APIs y software de minería de datos.
- Identificación de las plataformas y sitios que se encargan de la verificación de contenido político en Colombia y evaluar su eficacia para divulgar a los usuarios.
- Revisión de la literatura sobre el impacto de la desinformación y las noticias falsas en la sociedad colombiana, las estrategias aplicadas y las tecnologías empleadas, para establecer una base teórica sólida para el proyecto.
- Identificación de herramientas y técnicas para el monitoreo de contenidos verificados y no verificados en Twitter.
- Definición de indicadores de veracidad y divulgación de las noticias políticas en Twitter para la validación del prototipo desarrollado.
- Establecimiento de procedimientos para la validación del prototipo, considerando aspectos como la precisión y efectividad del mismo en la prevención de la desinformación y propagación de noticias falsas.

La importancia de esta fase radicó en establecer las bases teóricas y técnicas para el desarrollo del prototipo y su posterior validación.

1.6.2. Diseño y Análisis:

Las siguientes actividades permitieron definir los elementos del prototipo del sistema de alerta temprana para la detección y prevención de noticias falsas:

- Definición de las métricas y métodos para evaluar la eficacia del modelo.
- Diseño del prototipo del sistema de recolección y almacenamiento de datos que permita la actualización constante de la base de datos de noticias verificadas.
- Definición de funcionalidades y servicios del prototipo del sistema de alerta temprana para los usuarios, considerando la usabilidad y la experiencia del usuario.
- Definición de los métodos y técnicas para realizar fact checking social, teniendo en cuenta la eficacia y eficiencia de la metodología y su adaptación a la realidad de Colombia.
- Desarrollo de un sistema de machine learning que identifique contenido falso de forma no supervisada, considerando la disponibilidad y calidad de los datos en Twitter.
- Diseño de estrategias educativas que permitan a los usuarios identificar el contenido falso y evitar la propagación de noticias falsas, teniendo en cuenta las necesidades y características de los diferentes grupos de usuarios.
- Diseño de las interfaces del prototipo del sistema que tendrán las diferentes funcionalidades, teniendo en cuenta la usabilidad y la experiencia del usuario.
- Desarrollo del sistema de fact checking social, que permita validar la veracidad de las noticias y dar retroalimentación a los usuarios.
- Desarrollo de los métodos y sistemas de educación para la prevención de noticias falsas, con el fin de capacitar a los usuarios para detectar y evitar la propagación de noticias falsas.
- Definición del método de consumo de noticias verificadas de otras plataformas.

Todo esto facilitó el desarrollo del sistema de fact-checking social y la prevención de noticias falsas en Twitter, ofreciendo información verificada y fiable a los usuarios.

1.6.3. Desarrollo:

Durante esta fase, se emplearon técnicas de recolección y análisis de datos para evaluar la veracidad de las noticias, y se construyó un prototipo que permitió monitorear el contenido verificado y no verificado a través de las interacciones de los usuarios.

Las actividades desarrolladas fueron:

- Desarrollo del sistema de recolección y análisis de noticias de la API de Twitter.
- Desarrollo del prototipo de monitoreo de contenido verificado y no verificado mediante la integración de las interacciones y posturas de los usuarios.
- Implementación del sistema de clasificación de noticias y verificar su efectividad en la detección de noticias falsas.
- Integración del modelo de machine learning para la detección de noticias falsas con el sistema de clasificación de noticias.
- Implementación del servicio de notificación para alertar a los usuarios sobre noticias falsas detectadas.

1.6.4. Validación:

Para evaluar la efectividad del prototipo del sistema, durante esta fase, se llevaron a cabo las siguientes actividades:

- Pruebas de funcionalidad y rendimiento en el prototipo del sistema.
- Identificación de errores
- Pruebas de usuario y evaluación de usabilidad y facilidad de uso.
- Realizar análisis de los datos recolectados para identificar patrones de propagación de noticias falsas y posibles perjuicios a la sociedad.
- Evaluar el impacto del prototipo del sistema en la prevención de la desinformación y propagación de noticias falsas.
- Realizar un análisis de los efectos del prototipo del sistema en la prevención de la desinformación y propagación de noticias falsas en la sociedad.
- Realizar un análisis de los resultados obtenidos en la validación del prototipo y establecer medidas de mejora para la próxima iteración del prototipo del sistema.
- Realizar ajustes y mejoras en la aplicación en función de los resultados obtenidos durante las pruebas y evaluaciones.

1.7. Aportes de la investigación

En el contexto actual de una sociedad altamente digitalizada y conectada, la proliferación de noticias falsas representa un desafío significativo para la información veraz y la toma de decisiones informadas. En esta sección, se presentan los aportes de la investigación, resaltando cómo el enfoque contribuye de manera significativa a mitigar la propagación de información engañosa y a preservar la integridad de la información en el entorno digital.

- **Reducción de la propagación de noticias falsas:** El prototipo del sistema de alerta temprana desarrollado desempeña un papel crítico en la validación de la veracidad del contenido informativo, evitando de esta manera que los usuarios interactúen con información engañosa. Este enfoque proactivo tiene como consecuencia la mitigación del impacto negativo que pueden tener las noticias falsas y la desinformación en la opinión pública.
Al detectar y notificar a los usuarios de manera temprana acerca de la autenticidad del contenido, el sistema proporciona una herramienta esencial para promover la toma de decisiones informadas y la confianza en la información que se consume en el entorno digital.
- **Veracidad de la información:** En ciertas circunstancias, organizaciones radicales pueden utilizar noticias falsas como un medio para difundir su propaganda y sembrar el temor. La implementación de un sistema de alerta temprana se presenta como una medida eficaz destinada a detectar contenido engañoso y prevenir actividades potencialmente peligrosas. Esta iniciativa no solo contribuye a reforzar la seguridad pública, sino también a intensificar la lucha contra el extremismo
- **Confianza en la información consumida:** Mediante la verificación de la autenticidad y la disminución de la difusión de noticias falsas, los ciudadanos pueden confiar en la información que reciben, lo que a su vez fortalecerá su participación cívica y su capacidad para tomar decisiones informadas de manera más efectiva.
- **Protección de la democracia:** Las noticias falsas pueden ser empleadas como instrumentos de desestabilización en los procesos democráticos y en el sistema político en su conjunto. La implementación de un sistema de alerta temprana se convierte en un valioso aliado en la protección de la integridad del proceso democrático, al detectar y, en la medida de lo posible, prevenir la difusión de información engañosa.
- **Educación:** Educar a las personas para que sean críticas con la información que reciben y puedan discernir entre lo verdadero y lo falso.

1.8. Estructura del documento

El documento a presentar se estructura en seis capítulos, cada uno de los cuales se describe a continuación:

- **Capítulo 1:** En este capítulo se expone el planteamiento del problema que constituye el núcleo del proyecto de investigación, se detallan los objetivos propuestos, se describe la metodología empleada y se presenta la estructura del documento en su totalidad.
- **Capítulo 2:** En este capítulo, se expone el marco teórico, se analizan los trabajos relacionados en el campo y se identifican las brechas que aún persisten en relación al trabajo de investigación presentado.
- **Capítulo 3:** En este capítulo, se expone el proceso de planificación, ejecución y análisis de la información.
- **Capítulo 4:** En este capítulo, se presenta el desarrollo del prototipo funcional derivado de la investigación.
- **Capítulo 5:** En este capítulo, se exponen las pruebas de validación del prototipo funcional, junto con el análisis de los resultados.
- **Capítulo 6:** En este capítulo, se presentan las conclusiones derivadas de la investigación, además de posibles trabajos que podrían llevarse a cabo.

La figura 1.2 muestra la estructura del documento:

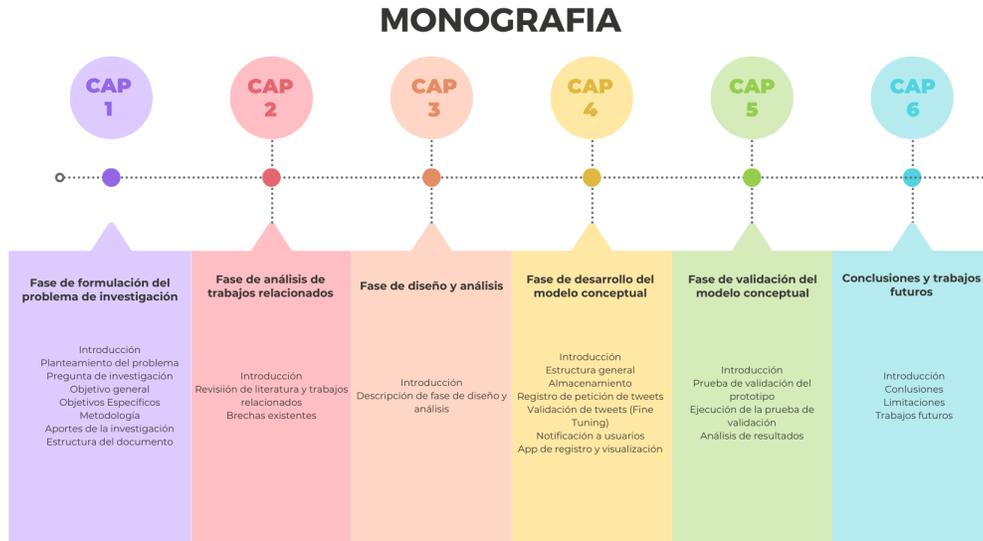


Figura 1.2: Estructura de la monografía (Figura propia)

CAPÍTULO 2

FASE DE ANÁLISIS DE TRABAJOS RELACIONADOS



2.1. Introducción

En este capítulo, se expone el marco teórico, se analizan los trabajos relacionados en el campo y se identifican las brechas que aún persisten en relación al trabajo de investigación presentado.

De este modo, la Figura 2.1, presenta el contenido abordado en el Capítulo 2:

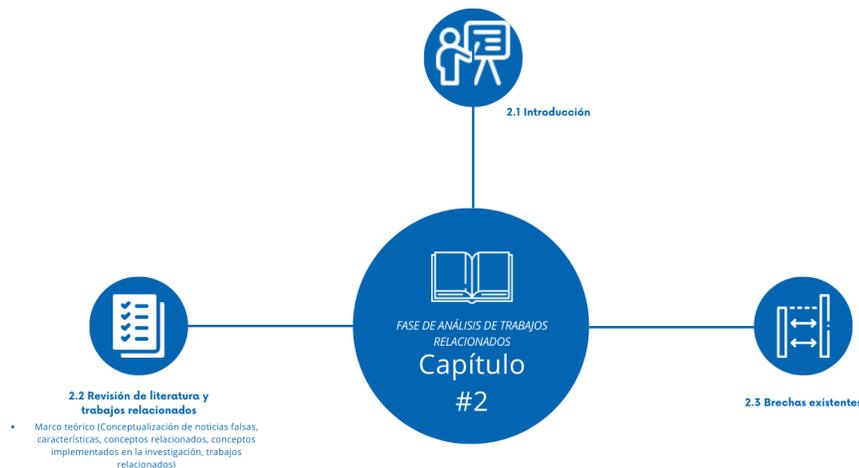


Figura 2.1: Mapa de contenido capítulo 2 (Figura propia)

2.2. Revisión de literatura y trabajos relacionados

2.2.1. Marco teórico

Las noticias falsas conocidas como "fake news", han emergido como un fenómeno creciente y preocupante en la era de la información digital [4]. Este fenómeno se caracteriza por la difusión deliberada de información errónea o engañosa, a menudo presentada como "noticias legítimas", con el propósito de influir en la opinión pública, desinformar o generar confusión [5].

El estudio de las fake news ha atraído la atención de académicos, investigadores y profesionales de diversos campos, ya que plantea desafíos significativos para la democracia, la confiabilidad de la información y la toma de decisiones informadas.

En este marco teórico, se abordará la conceptualización, las dimensiones y las implicaciones de las fake news, basándose en la literatura académica disponible.

Conceptualización de noticias falsas:

Las noticias se entienden como metainformación y pueden incluir lo siguiente: [15]

- **Fuente:** Editores de noticias, autores, enlaces a sitios web y redes sociales [3].
- **Titular:** Presentación del tema central de la noticia con un breve texto para captar la audiencia [3].
- **Cuerpo de noticia:** Descripción detallada de la noticia, incluyendo aspectos destacados y características del editor [3].
- **Contenido multimedia:** Fragmento del cuerpo informativo que ofrece una representación visual con el propósito de facilitar la comprensión del contenido [3].
- **Enlaces:** Enlaces a otras fuentes de noticias [3].

La expresión “fake news” fue seleccionada como la palabra del año por el Diccionario Macquarie en el año 2016. Las noticias falsas han atraído considerable atención por parte de la comunidad investigadora, dando lugar a diversas interpretaciones y perspectivas. En su obra [15], los autores han definido las noticias falsas como “información noticiosa que es deliberadamente falsa y verificable como tal”. Por su parte, Alcott y Gentzkow [4] han ofrecido una definición más rigurosa de las noticias falsas, considerándolas como “artículos informativos intencionalmente falsos que tienen la capacidad de inducir al engaño en los lectores”.

Otra interpretación, [16], concibe las noticias falsas como “información manufacturada que emula la estructura de las noticias mediáticas, pero que difiere en cuanto a su proceso de generación y propósito”. Además, en su enfoque [17], los autores han abordado las noticias falsas en diversas manifestaciones, incluyendo noticias falsas, engañosas o creativas, examinando múltiples atributos y características de la información difundida. En [27], se presenta una amplia definición de noticias falsas como “información noticiosa inauténtica”, junto con una definición más restringida que se refiere a las noticias falsas como “información noticiosa intencionalmente falsa publicada por medios de comunicación”.

Estas definiciones han sido utilizadas en enfoques previos para la detección de noticias falsas y ofrecen una base sólida para el estudio y análisis de este fenómeno en constante evolución [18],[3],[17],[19].

Características de las noticias falsas:

Las noticias falsas, han sido objeto de diversas interpretaciones debido a la diversidad de contextos en los que se difunden.

Estas noticias se caracterizan por compartir una serie de atributos comunes, los cuales incluyen:

- **Efecto cámara de eco:** Las cámaras de eco [20], en términos generales, pueden definirse como entornos que se enfocan en las opiniones de usuarios que comparten la misma orientación política o creencias sobre un determinado tema. Estas opiniones se refuerzan a través de interacciones repetidas con otros usuarios que poseen inclinaciones y actitudes similares. La credibilidad social [21] y la heurística de recurrencia [21] (es decir, la tendencia a buscar información que concuerde con las opiniones preexistentes) pueden ser la causa de la aparición de cámaras de eco en las redes sociales [15],[22], [23],[24].
Cuando las noticias carecen de información, la credibilidad social puede emplearse como criterio para evaluar su veracidad. No obstante, muchas personas aún las perciben como creíbles y las comparten, lo que conduce a la aceptación generalizada de tales noticias como verídicas. Una heurística de recurrencia se forma cuando las personas se encuentran con noticias periódicamente, lo que naturalmente lleva a la aprobación de la información, incluso si se trata de noticias falsas.
- **Intención de engañar:** Esta característica se evidencia a partir de la premisa de que “nadie accidentalmente genera información inexacta con la forma típica de los reportajes noticiosos, y el género de las noticias falsas se concibe intencionadamente con el propósito de inducir al “engaño”[16].
Los motivos subyacentes al engaño pueden ser tanto de índole política/ideológica como financiera [4],[5],[25],[26].
Además, también pueden surgir situaciones en las que las noticias falsas se difundan con la finalidad de brindar entretenimiento, diversión o, como se plantea en [27], “generar controversia”.
- **Cuenta maliciosa:** En la actualidad las noticias en las plataformas de redes sociales provienen tanto de usuarios legítimos como de perfiles ficticios. Aunque es importante destacar que la propagación de noticias falsas se origina principalmente en cuentas no auténticas, es relevante subrayar que existen individuos reales que, de igual manera, participan en la difusión de desinformación. Las cuentas dedicadas principalmente a la diseminación de noticias falsas son comúnmente denominadas como “cuentas maliciosas”[27]. Estas cuentas maliciosas se dividen en tres categorías principales: bots sociales, trolls y usuarios de cíborgs [15].

Los bots sociales representan perfiles en redes sociales controlados por algoritmos informáticos. Se considera que una cuenta es maliciosa cuando su principal finalidad es divulgar información perjudicial y desempeña un papel preponderante en la creación y propagación de noticias falsas [28]. Estas cuentas maliciosas también pueden automatizar la publicación de noticias y mantener interacciones con otros usuarios de las redes sociales.

Los trolls, por su parte, son individuos reales que interfieren en las comunidades en línea con la intención de suscitar respuestas emocionales de los usuarios de las redes sociales [15]. Los trolls buscan deliberadamente manipular la información con el propósito de modificar las opiniones de los demás [28], generando emociones negativas entre los usuarios y, como consecuencia, sembrando dudas y desconfianza [15]. Esta confusión progresiva deja a los usuarios en un estado de incertidumbre, dificultando su capacidad para discernir entre información auténtica y falsa. Con el tiempo, esto puede llevar a que los usuarios abriguen dudas sobre la veracidad de la información y acepten creencias erróneas y datos falsos.

Por último, los usuarios de cíborgs son cuentas maliciosas creadas por individuos reales pero gestionadas en gran medida a través de programas y algoritmos. De este modo, los cíborgs demuestran ser altamente eficaces en la difusión de noticias falsas [15].

- **Autenticidad:** Esta función se enfoca en determinar la imparcialidad de una noticia [27]. Las afirmaciones factuales pueden ser corroboradas como verídicas o falsas. En contraste, las opiniones subjetivas no se clasifican como afirmaciones factuales; únicamente las opiniones objetivas son consideradas como tales. Las declaraciones de hecho nunca pueden ser incorrectas, ya que una declaración que puede ser refutada no se clasifica como una afirmación de hecho [29]. Las declaraciones no fácticas son afirmaciones sobre las cuales podemos tener diferentes puntos de vista. En otras palabras, la precisión de una noticia puede variar, oscilando entre incorrecta, correcta o completamente inexacta. En su mayoría, las noticias falsas suelen incluir afirmaciones que carecen de objetividad.

Conceptos relacionados con las noticias falsas:

Diversos enfoques teóricos han abordado el fenómeno de las noticias falsas, a través de la aplicación de las características inherentes a este tipo de información, es posible redefinir dichos conceptos con el propósito de establecer claras distinciones entre ellos.

- **Las noticias falsas** [30], [31], son noticias que contienen declaraciones no objetivas de cuentas maliciosas que pueden causar el efecto de cámara de resonancia con intenciones indefinidas.

- **La desinformación** [32], son noticias o no noticias que contienen declaraciones no objetivas de cuentas maliciosas que pueden causar el efecto de cámara de eco, con la intención de engañar al público.
- **La selección selectiva** [33], son noticias o no noticias que contienen declaraciones fácticas comunes de cuentas maliciosas y pueden causar el efecto de cámara de eco, con la intención de engañar al público.
- **Rumor** [34], son noticias o no noticias que contienen declaraciones fácticas o no fácticas de cuentas maliciosas y pueden causar el efecto de cámara de resonancia con intenciones indefinidas.
- **La información falsa** [35], es noticia o no noticia de declaraciones no fácticas de cuentas maliciosas que pueden provocar el efecto de cámara de eco, con la intención de engañar al público.
- **La manipulación** [36], son noticias en los mercados que contienen declaraciones no objetivas de cuentas maliciosas que pueden causar el efecto de cámara de eco, con la intención de engañar al público.
- **Las noticias engañosas** [4], [15], son noticias que contienen declaraciones no objetivas de cuentas maliciosas que pueden causar el efecto de cámara de eco, con la intención de engañar al público.
- **La desinformación** [23], son noticias o no noticias que contienen declaraciones no objetivas de cuentas maliciosas que pueden causar el efecto de cámara de eco con intenciones indefinidas.
- **Los hechos falsos** [37], son información indefinida (noticia o no) que comprende declaraciones no fácticas de cuentas maliciosas que pueden causar el efecto de cámara de resonancia, con la intención de engañar al público.
- **La propaganda** [35], es información sesgada (noticia o no) que comprende declaraciones indefinidas (fácticas o no fácticas) sobre eventos principalmente políticos provenientes de cuentas maliciosas y que pueden causar el efecto de cámara de eco, con la intención de engañar al público.

Conceptos implementados en la investigación:

- **Efectividad:** Según la Real Academia Española (RAE), el concepto efectividad se define como: “Capacidad de lograr el efecto que se desea o se espera” [13].

En otras palabras, la efectividad se relaciona con la capacidad de lograr los resultados deseados de manera satisfactoria, teniendo en cuenta la eficacia

en la consecución de objetivos y la eficiencia en la utilización de recursos, y es un término ampliamente discutido y analizado en la literatura académica y profesional.

- **Contexto:** La definición de contexto según la RAE es un “entorno físico o de situación, político, histórico, cultural o de cualquier otra índole, en el que se considera un hecho” [28], esto se refiere a la situación en que sucede algún acontecimiento y estas particularidades pueden incluir el lugar en donde sucede, el tiempo, el entorno y las personas implicadas.

El contexto es bien definido y además cambia por los participantes en una situación específica, en [29], se refieren cuatro clases de contextos: el primero es el contexto situacional en el cual se hace referencia al entorno físico en donde se encuentra u ocurre la interacción; el segundo, es el contexto sociocultural donde, se hace referencia al “conocimiento de fondo”, es decir, a la información que viene dada por la sociedad y la cultura dentro del entorno en que se encuentre; el tercero es el contexto cognitivo que se hace referencia a “inferencias y otras formas de razonamiento”, en este tipo se tiene en cuenta la noción propia que se tenga al momento de realizar la interacción y de la cual sacamos primeras impresiones del contexto y, por último, es el contexto lingüístico en el que se hace alusión a la “información lingüística que sigue a un enunciado”, esto quiere decir que el contexto depende, por ejemplo, de lo que se haya hablado en el momento de interacción.

El contexto de uso de un producto tiene que ver con las condiciones o el lugar en donde se va a ser utilizado [38] y es importante identificar este contexto para obtener una usabilidad [39] aceptable que brinde una buena experiencia al usuario con la finalidad de conseguir los objetivos propuestos con dicho producto.

- **Prototipo:** En [38], se define el concepto de prototipo como “una representación física o digital de elementos críticos del diseño y una herramienta iterativa para mejorar la comunicación, permitir el aprendizaje y fundamentar la toma de decisiones en cualquier punto del proceso de diseño”. Un prototipo se puede utilizar para probar y evaluar objetivos o ideas en una primera instancia antes de tener un producto final, es importante para ayudar a entender de mejor manera el cumplimiento de los requerimientos o necesidades de los usuarios para mejorar en los puntos débiles del diseño o del funcionamiento. Los prototipos ayudan a responder a preguntas relacionadas con la deseabilidad, la viabilidad y la factibilidad del proyecto en cuestión [40].

Un prototipo [41] puede ser tanto un modelo en tres dimensiones como de igual forma modelos no físicos como bocetos o representaciones digitales;

se tiene la particularidad que los prototipos se pueden realizar ágilmente con presupuesto limitado para ayudar a identificar problemas y retroalimentar su implementación y construcción [32]. La elección del prototipo a utilizar depende de lo que se quiera conseguir en el proyecto teniendo en cuenta además el factor de tiempo y dinero que se pueda implementar en el desarrollo del mismo, en [44] se da una representación para la construcción de un prototipo y se mencionan tres pasos principales:

1. La elaboración de un patrón, modelo o propuesta.
2. Revisión por parte de los usuarios.
3. Se pregunta si cumple con lo esperado, a partir de este momento si hay observaciones se vuelve a una etapa de diseño y se vuelve al mismo ciclo, cuando no existan observaciones se evalúa si se alcanzaron los objetivos.

En la Figura 2.2 se observa la estructura de un prototipo:

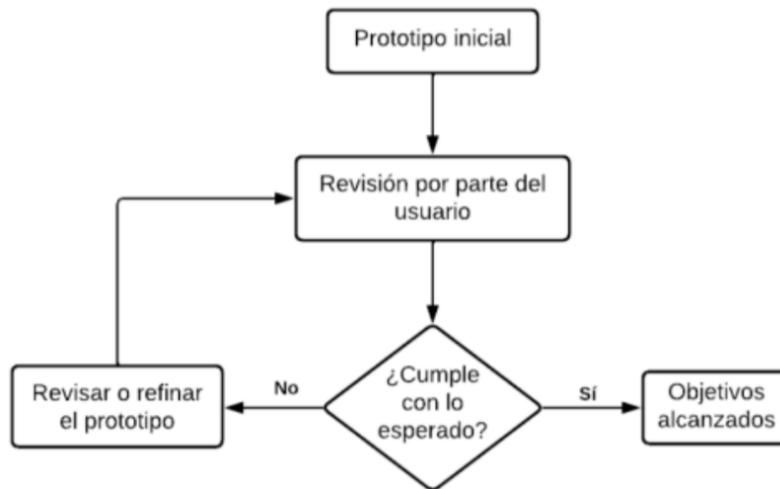


Figura 2.2: Estructura del prototipo (Figura tomada de 41)

Con un prototipo se puede dar validez a un proyecto por la capacidad de probar, demostrar y mejorar la idea o el planteamiento por medio de este, para garantizar que el objetivo propuesto se alcanzó y que el prototipo es válido o exitoso, es importante que tanto el diseñador como el usuario deban estar implicados en la parte del proceso de diseño [41].

- **Caracterización de usuarios:** [42] Según la Real Academia Española (RAE) el concepto caracterizar se define como: “Determinar los atributos peculiares de alguien o de algo, de modo que claramente se distinga de los demás” [26]. Entonces, la caracterización [43] puede ayudar a definir cómo es la descripción de un individuo basada en características propias de este (la apariencia, su forma de ser, la manera en que se expresa, etc.). Al intentar caracterizar una persona se pueden tener diferencias en los puntos de vista debido a que esto se realiza de una manera subjetiva y depende completamente de las opiniones y percepciones de la persona que realice la caracterización.

En una investigación es importante definir muy bien las características del individuo o del objeto a estudiar, para esto se deben establecer los métodos, las técnicas y los instrumentos que se utilizan para la recolección de la información [27], con la finalidad de comprender de mejor manera con quien se está trabajando ya que esto puede ayudar a acortar el estudio y definir unas métricas claras a seguir para obtener los resultados más apropiados al tener cómo referencia dicha caracterización previamente realizada.

- **Encuesta CSUQ:** La encuesta CSUQ es una herramienta diseñada para evaluar la usabilidad [44] de sistemas informáticos y aplicaciones. Desarrollada por Lund y Wixon en 1985, la CSUQ se ha convertido en un instrumento ampliamente utilizado en la industria y la investigación para medir la satisfacción del usuario con respecto a la interfaz, la facilidad de uso y la eficiencia de un sistema computacional [45].

2.2.2. Trabajos relacionados

Con el objetivo de entender las tendencias actuales en la investigación sobre la identificación y propagación de noticias falsas, se realizó un estado del arte utilizando la metodología del mapeo sistemático de Kitchenham [46] (2010). En el proceso de evaluación de los artículos recopilados en este mapeo sistemático, se priorizó el análisis de los resúmenes y las conclusiones de cada estudio para determinar su relevancia en el contexto de la investigación sobre noticias falsas. Como resultado de este riguroso proceso, se identificaron 12 artículos de investigación que destacan significativamente en la comprensión y el abordaje de la problemática de las noticias falsas.

Estos artículos ofrecen una perspectiva diversa y detallada sobre una serie de aspectos cruciales relacionados con la detección y mitigación de las noticias falsas en el entorno digital actual. A continuación, se destacan los temas centrales abordados por cada uno de estos estudios:

1. Impacto de las Plataformas Digitales y Verificadores de Información:

El primer artículo, “El papel de las plataformas digitales y los verificadores de información en la difusión de las fake news” [8], se enfoca en analizar el impacto de las noticias falsas en la política y las redes sociales en múltiples países, considerando el papel de las plataformas digitales y los verificadores de información. Propone investigar cómo las noticias falsas, la política y las redes sociales están cambiando el debate público en diversos países, utilizando ejemplos específicos como las elecciones presidenciales de Estados Unidos en 2016, el referéndum del Brexit en el Reino Unido y la campaña de desinformación en Colombia. La investigación se enfoca en el impacto de las noticias falsas en la intención de voto, la forma de hacer periodismo y la credibilidad de los medios. Además, se busca encontrar posibles soluciones para frenar el poder de las noticias falsas en este panorama de intoxicación a través de entrevistas con referentes del periodismo colombiano.

2. Detección de Noticias Falsas en Twitter: El segundo artículo, “Modelo para la detección de noticias falsas en formato texto en la red social Twitter, aplicado al contexto político colombiano de las elecciones presidenciales de 2022:” [1], propone una solución para abordar la ausencia de modelos de detección de noticias falsas en el contexto político colombiano en la red social Twitter. La propuesta se basa en la metodología CRISP-DM, que consta de seis fases, y se utiliza una serie de modelos, técnicas y herramientas de la ciencia de datos para abordar la solución del problema. Estos modelos incluyen Random Forest, Naive Bayes, Support Vector Classifier, Regresión Logística, XG-Boost, redes neuronales tradicionales y modelos basados en atención como BERT. Los resultados obtenidos muestran que el modelo basado en BERT logró mejores resultados en comparación con los modelos tradicionales. Además, la validación de la propuesta se realizó mediante varios experimentos con diferentes corpus de noticias, lo que permitió obtener un modelo de analítica capaz de detectar noticias falsas en el contexto político colombiano con un buen nivel de desempeño.

3. Marco de Trabajo para la Detección de Noticias Falsas Extremas: “EFND:

A semantic visual, and socially augmented deep framework for extreme fake news detection ” [47], propone un nuevo marco de trabajo (framework) para la detección de noticias falsas extremas, llamado EFND. Este marco de trabajo utiliza una combinación de técnicas de procesamiento de lenguaje natural (PLN), visión por computadora y análisis de redes sociales para mejorar la precisión de la detección de noticias falsas extremas.

El marco EFND se divide en cuatro etapas principales:

- a) **Extracción de características:** se utilizan técnicas de PLN para extraer características relevantes de los textos de las noticias, como el contenido temático, la polaridad y la estructura gramatical.
- b) **Análisis visual:** se utilizan técnicas de visión por computadora para analizar el contenido visual de las noticias, como imágenes y videos, y extraer características relevantes, como la calidad de la imagen y la detección de manipulaciones visuales.
- c) **Análisis social:** se utilizan técnicas de análisis de redes sociales para analizar la estructura de la red social en torno a la noticia, como la propagación de la noticia y la participación de los usuarios en la discusión.
- d) **Clasificación:** se utiliza un modelo de aprendizaje profundo (deep learning) para clasificar las noticias como verdaderas o falsas.

El marco EFND se probó en un conjunto de datos de noticias falsas extremas y se comparó con otros enfoques existentes. Los resultados mostraron que EFND tuvo un rendimiento significativamente mejor en la detección de noticias falsas extremas en comparación con otros enfoques existentes.

4. **El Papel de las Advertencias en la Reducción de la Propagación de Noticias Falsas:** “Can warnings curb the spread of fake news ? The interplay warning, trust and confirmation bias” [48], examina cómo los usuarios de las redes sociales procesan las advertencias y cómo se ven afectados por factores como la confianza en la fuente, la confirmación de sus propias creencias y la influencia de la opinión de los demás. Los autores sugieren que, si se implementan de manera efectiva, las advertencias pueden ser una herramienta útil para combatir la propagación de noticias falsas y aumentar la responsabilidad de los usuarios al momento de compartir información.
5. **Detección de Manipulación de Plataformas y Propaganda Computacional en Twitter:** El quinto estudio “A framework to detect twitter platform manipulation and computational propaganda” [49], propone un marco teórico para detectar la manipulación de plataformas y la propaganda computacional en la plataforma de Twitter. Argumenta que la detección de la manipulación de plataformas y la propaganda computacional en Twitter es crucial para garantizar la integridad y la autenticidad de las conversaciones públicas en línea. El marco propuesto por los autores incluye la identificación de los actores y los comportamientos asociados con la manipulación de plataformas y la propaganda computacional, la medición de la magnitud de la manipulación y la propagación de la propaganda en la plataforma, y la evaluación de las consecuencias de la manipulación y la propaganda en la

opinión pública.

Los autores indican que su marco teórico puede ser utilizado por investigadores, responsables políticos y reguladores para detectar la manipulación de plataformas y la propaganda computacional en Twitter y, en última instancia, proteger la integridad y la autenticidad de las conversaciones públicas en línea.

6. **Detección de Noticias Falsas Basada en Opiniones Subjetivas:** El sexto artículo, “Fake news Detection Based in Subjective Opinions” [6], presenta un enfoque innovador basado en opiniones subjetivas para la detección de noticias falsas, considerando la subjetividad en la evaluación de la veracidad de las noticias. El enfoque propuesto por los autores tiene en cuenta la subjetividad inherente a la evaluación de la veracidad de las noticias, y utiliza técnicas de análisis de sentimientos para extraer opiniones subjetivas de los textos asociados con las noticias. Estas opiniones se utilizan para alimentar un modelo de clasificación que puede identificar noticias falsas con una alta precisión.

Argumentan que su enfoque puede ser utilizado por los usuarios de las redes sociales para evaluar la veracidad de las noticias en tiempo real, y también puede ser utilizado por las organizaciones de noticias y los reguladores para detectar y eliminar noticias falsas en línea.

7. **Análisis de Texto para la Investigación de Desinformación en Twitter:** El séptimo estudio, “Text analysis methods for misinformation-related research on Finnish language Twitter” [50], describe diferentes enfoques para el análisis de texto, incluyendo el análisis de sentimiento, el análisis de tópicos y la detección de eventos. También se presentan técnicas específicas para identificar la desinformación, como el análisis de la veracidad de los tweets y la detección de bots y cuentas falsas; propone que los métodos expuestos pueden ser utilizados para investigar la desinformación en Twitter en el idioma finlandés y proporciona una guía para la selección de técnicas y herramientas apropiadas para este propósito.

8. **Revisión de Técnicas y Enfoques para la Detección de Noticias Falsas en Twitter:** El octavo artículo, “Fake News detection on Twitter” [51], expone una revisión de las técnicas y enfoques utilizados para detectar noticias falsas en Twitter. En él, se discuten las características de las noticias falsas, los desafíos asociados con su detección en la plataforma de Twitter y las técnicas de detección de noticias falsas existentes.

Las técnicas de detección son:

- Enfoques basados en características lingüísticas y de contenido, como el análisis de sentimientos, la detección de polaridad y la clasificación de texto.

- Enfoques basados en la verificación de hechos y la autenticidad, como el seguimiento de fuentes y la verificación de imágenes y videos.
 - Enfoques basados en el aprendizaje automático, como la clasificación de textos con algoritmos de aprendizaje supervisado y no supervisado.
9. **Impacto de las Noticias Falsas en el Comportamiento de Voto:** El noveno artículo, “¿Does fake news affect voting behaviour” [7], analiza si las noticias falsas tienen un impacto en el comportamiento de voto de las personas. Los autores se basan en datos de una encuesta realizada en el Reino Unido después del referéndum del Brexit en 2016, donde se les preguntó a los encuestados sobre su exposición a noticias falsas y su comportamiento de voto.
10. **Enfoque Híbrido en la Detección de Noticias Falsas en Tiempo Real:** El décimo estudio, “Hybrid Approach and Architecture to detect Fake News on Twitter ,in Real-Time using Neural Networks” [52], propone un enfoque híbrido y una arquitectura para detectar noticias falsas en Twitter en tiempo real utilizando redes neuronales. La propuesta combina técnicas de procesamiento de lenguaje natural, aprendizaje automático y análisis de redes sociales para identificar noticias falsas en Twitter de manera eficiente y precisa.

La arquitectura propuesta consta de varias capas, que incluyen una capa de preprocesamiento, una capa de extracción de características, una capa de clasificación y una capa de retroalimentación. Se utilizan técnicas de procesamiento de lenguaje natural como tokenización, eliminación de stopwords y lematización para preprocesar los datos. Luego, se utilizan algoritmos de aprendizaje automático, como redes neuronales convolucionales (CNN) y redes neuronales recurrentes (RNN), para extraer características de los datos y clasificarlos en noticias verdaderas o falsas.

Además, el enfoque híbrido propuesto también utiliza técnicas de análisis de redes sociales para analizar la propagación de noticias en Twitter y detectar patrones de difusión de noticias falsas.

11. **Comparación de Modelos de Aprendizaje Automático:** El artículo, “Identifying Fake News with various machine learning model” [53], propone identificar noticias falsas utilizando varios modelos de aprendizaje automático. El objetivo del estudio es comparar el rendimiento de diferentes modelos de aprendizaje automático para la detección de noticias falsas y determinar cuál es el modelo más efectivo. Para lograr esto, los autores recopilaron datos de noticias tanto verdaderas como falsas y crearon un conjunto de datos de entrenamiento y prueba. Luego, utilizaron varios modelos de aprendizaje automático, incluyendo Naive Bayes, Regresión Logística, SVM (Support

Vector Machines), Random Forest y Redes Neuronales, para identificar noticias falsas.

Los autores también utilizaron técnicas de preprocesamiento de lenguaje natural para preparar los datos antes de entrenar los modelos, como la eliminación de stopwords, la tokenización y la lematización. Además, utilizaron técnicas de extracción de características para extraer información relevante de los datos, como el conteo de palabras y la frecuencia de términos.

Los resultados mostraron que los modelos de Redes Neuronales y SVM tuvieron un mejor rendimiento en la detección de noticias falsas en comparación con otros modelos. Además, los autores encontraron que la combinación de diferentes características, como la frecuencia de términos y el conteo de palabras, mejoró el rendimiento de los modelos.

En conclusión, el artículo propone utilizar varios modelos de aprendizaje automático para la detección de noticias falsas y comparar su rendimiento. Los autores utilizaron técnicas de preprocesamiento de lenguaje natural y extracción de características para preparar los datos y entrenar los modelos. Los resultados muestran que los modelos de Redes Neuronales y SVM son los más efectivos para la detección de noticias falsas.

12. **Análisis de Sentimiento de Datos de Twitter:** El último artículo, “Deep sentiment analysis od twitter data using a Nybrid Ghost convolution Neural Network Model” [2], propone un modelo híbrido de red neuronal para realizar un análisis profundo de sentimiento en los datos de Twitter. Además, presenta los resultados de experimentos realizados con diferentes conjuntos de datos de Twitter y muestra que el modelo propuesto supera a otros modelos de referencia en términos de precisión de clasificación y eficiencia computacional. El modelo también puede ser útil para analizar grandes volúmenes de datos de redes sociales y proporcionar información valiosa para la toma de decisiones en diversos campos, como el marketing, la opinión pública y la investigación de mercado.

En resumen, estos trabajos relacionados brindan una visión integral de las tendencias de investigación actuales sobre noticias falsas, destacando diversos enfoques y estrategias para abordar este desafío en constante evolución en el entorno digital. Cada artículo contribuye de manera única a la comprensión y el trabajo alrededor de la propagación de noticias falsas.

2.3. Brechas existentes

Como consecuencia de la revisión de la literatura, se detectaron áreas de mejora en el ámbito de la detección y prevención de noticias falsas.

En ese sentido, se identificaron algunas brechas que surgen desde la perspectiva de los siguientes criterios:

- **Utilización de métodos de verificación y detección de noticias falsas:** El sistema debe demostrar la implementación de técnicas y herramientas efectivas para identificar y verificar la veracidad de las noticias, asegurando la detección de contenido falso o engañoso.
- **Utilización de las interacciones y posturas de los usuarios de las redes sociales:** El sistema debe aprovechar la información generada por los usuarios en las redes sociales para mejorar su capacidad de detección y prevención de noticias falsas, considerando las interacciones, opiniones y comportamientos de los usuarios como indicadores relevantes.
- **Implementación de estrategias para prevenir la propagación de noticias falsas en las redes sociales:** El sistema debe contar con medidas y técnicas específicas destinadas a mitigar la difusión de contenido falso en las redes sociales, estableciendo mecanismos efectivos para evitar su viralización y limitar su alcance.
- **Análisis de los efectos de la propagación de noticias falsas sobre la desinformación ciudadana:** El sistema debe realizar un análisis exhaustivo de los impactos y consecuencias que la propagación de noticias falsas tiene en la desinformación de la ciudadanía, proporcionando una comprensión clara de las implicaciones de este fenómeno.
- **Provisión de información clara y fácilmente comprensible para identificar noticias falsas:** El sistema debe brindar a los usuarios de las redes sociales información precisa, transparente y de fácil comprensión que les permita identificar y discernir entre noticias falsas y verídicas, promoviendo la alfabetización mediática.
- **Incorporación de alertas tempranas para notificar a los usuarios sobre la posible propagación de noticias falsas:** El sistema debe incluir mecanismos de alerta temprana que notifiquen a los usuarios de las redes sociales sobre la presencia y posible propagación de noticias falsas, con el objetivo de prevenir su difusión y minimizar su impacto.

Las brechas identificadas evidencian que existe investigación y análisis alrededor de la problemática de noticias falsas, con diversos métodos de detección para este tipo de contenido. Sin embargo, estos métodos no abordan mecanismos de

prevención y alerta ante las noticias falsas, careciendo de aplicaciones que disminuyan el efecto de la propagación de estas noticias.

Por lo tanto, es necesario indagar qué mecanismos previenen y alertan de forma práctica a los usuarios de las redes sociales, generando conciencia sobre las consecuencias de las noticias falsas y cómo se pueden detectar de forma temprana.

CAPÍTULO 3

FASE DE DISEÑO Y ANÁLISIS



3.1. Introducción

En el tercer capítulo se detalla el proceso de recolección y análisis de noticias. En este contexto, se establecen las páginas de las cuales se extrajo la información, se define el algoritmo a seguir, y se estructura el conjunto de datos necesario para llevar a cabo dicho análisis. Luego, se verificó el contenido recopilado, seguido de la creación de un dataset específico para el proceso de Fine-Tuning con el fin de refinar la calidad de los datos. A lo largo de esta sección, se expondrán los diversos métodos utilizados, desde la recopilación de noticias hasta el procesamiento de la información.

Finalmente, se presentan gráficos ilustrativos que ofrecen un panorama visual del comportamiento de la información recolectada, enriqueciendo así la comprensión de los resultados obtenidos.

De este modo, la Figura 3.1, presenta el contenido abordado en el Capítulo 3:

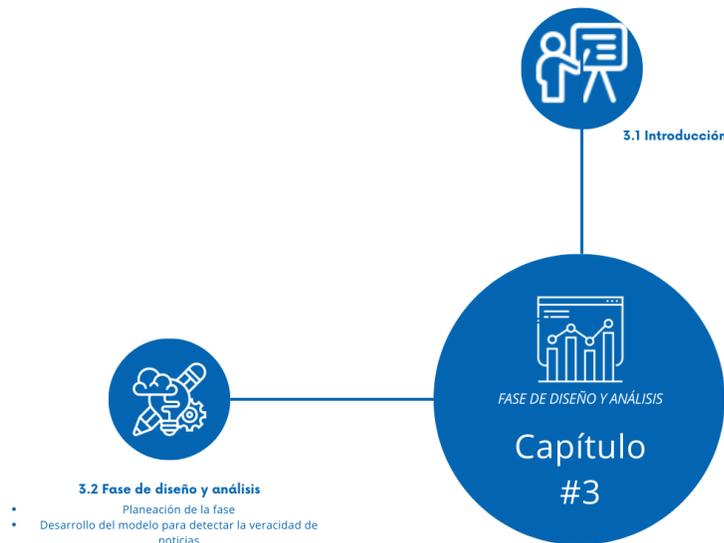


Figura 3.1: Mapa de contenido capítulo 3 (Figura propia)

3.2. Fase de diseño y análisis

3.2.1. Planeación de la fase

La creación del modelo de análisis de veracidad implica una serie de etapas en las cuales se abordan la recopilación de datos, la elección de noticias, el proceso de capacitación, la validación y la evaluación de resultados.

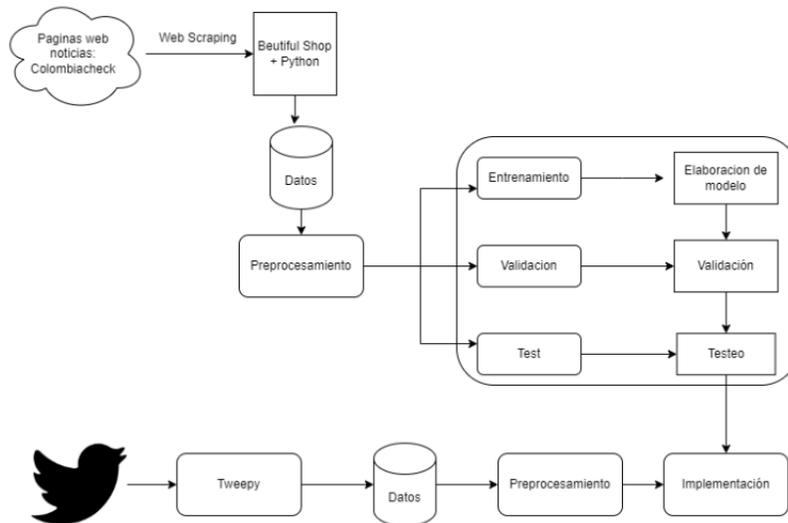


Figura 3.2: Modelo empleado para el desarrollo de la investigación (Figura propia)

La figura 3.2 ilustra el procedimiento empleado para obtener noticias verificadas por parte de los **FactCheckers**, que se logró mediante la extracción de datos de la página Colombia Check, utilizando técnicas de web scraping implementadas en Python.

Los datos recopilados permitieron la creación de un conjunto de noticias, seguido de un proceso de filtrado con el objetivo de conformar un dataset que se enfocó en la información verificada relacionada con el ámbito político de Colombia.

Este conjunto de datos se utilizó en el proceso de ajuste para entrenar un modelo de inteligencia artificial.

3.2.2. Desarrollo del modelo para detectar la veracidad de noticias

Definición de medios de selección de noticias:

Diversos canales de comunicación difunden volúmenes significativos de información diariamente. Sin embargo, una parte sustancial de esta información carece

de credibilidad y se comparte a través de plataformas en línea y redes sociales con diversos propósitos. Por lo tanto, se requiere el establecimiento de criterios rigurosos para discernir la veracidad de las noticias.

Estos desempeñan un papel fundamental en la selección de noticias para el análisis en el presente estudio. Se basan en múltiples atributos, que incluyen la fuente de la noticia, su consistencia con otras fuentes, la calidad del contenido, la reputación del emisor y la presencia de evidencia.

A continuación, se detallan:

1. **Fuente confiable:** Es imperativo realizar una evaluación de la fuente de información. Esta debe centrarse en la reputación y credibilidad de la entidad emisora, que podría tratarse de instituciones de prestigio, sitios web de renombre o especialistas altamente reconocidos en la disciplina correspondiente.
En general, las fuentes oficialmente reconocidas y consolidadas suelen ofrecer una mayor fiabilidad en contraposición a aquellas cuya procedencia es incierta o cuestionable.
2. **Contraste de múltiples fuentes:** Validar si la información cuenta con respaldo proveniente de diversas fuentes independientes y confiables es crucial. Sin embargo, si únicamente se encuentra respaldada por una única fuente sin apoyo adicional, podría ser motivo de precaución.
3. **Calidad y reputación del medio:** Evaluar la calidad y la reputación del medio que publica la noticia.
4. **Evidencia y datos verificables:** Proporcionar información respaldada por pruebas tangibles, como investigaciones, estudios científicos, estadísticas confiables o testimonios de expertos.
5. **Redacción y estilo:** Es importante estar atento al tono y estilo de redacción presentes en la noticia.
6. **Sesgo y verificación cruzada:** Evaluar si la noticia exhibe algún sesgo evidente o busca influir en la opinión pública. Además, es fundamental corroborar si otras fuentes confiables también abordan la misma noticia, con el fin de obtener una perspectiva equilibrada.
7. **Comprobación de hechos:** Emplear herramientas de verificación de hechos y consultar sitios web especializados en la detección de noticias falsas es fundamental para corroborar la información. Algunos ejemplos de estas herramientas son Snopes, FactCheck.org y Politifact.

Definición de sitios web:

Diversos portales de noticias y revistas digitalizadas desempeñan un papel crucial en la disseminación de información a una audiencia considerable. No obstante, al

considerar aspectos de investigación, se aboga por la elección de medios especializados en la verificación y validación de contenido, comúnmente reconocidos como "fake checkers". En consonancia con estos criterios previos, se optó por la plataforma "Colombia Check", la cual se dedica a la verificación de noticias y contenido relacionado con la política colombiana. Esta elección aseguró la recopilación de información verificada y de alta calidad.

Recolección de noticias de medios verificados:

La obtención de noticias desde la plataforma web de ColombiaCheck involucra el uso de la biblioteca Python BeautifulSoup para llevar a cabo la solicitud y la extracción de datos. El proceso inició mediante el análisis de la estructura de la página, con el propósito de comprender cómo la información está organizada en el código HTML. Introducción tiene como objetivo establecer de manera sistemática qué campos serán observados y capturados. Posteriormente, se extrajo la URL de la página Colombia Check, que dirige a los encabezados de las noticias a ser recopiladas. Una vez definidos los campos, se ejecuta la solicitud a la página, y los datos extraídos incluyen:

- **Título:** El título de la noticia describe el titular del acontecimiento relatando de forma concreta lo ocurrido.
- **Cuerpo:** Describe de forma sencilla lo ocurrido, narrando el suceso principal, los afectados entre otras características.
- **Fecha:** Fecha en la cual ocurre el acontecimiento.
- **Tipo:** Describe si la noticia es falsa o verdadera.
- **Fuente:** Fuente de la cual es tomada la noticia.
- **Link:** Enlace de origen de la noticia.

Filtrado y creación de dataset de noticias:

Con los datos recopilados, se llevó a cabo un proceso de filtrado con el propósito de seleccionar exclusivamente las noticias vinculadas a la esfera política. Este proceso implica la aplicación de criterios anteriormente tratados.

Las noticias que pasaron por el proceso de filtrado fueron almacenadas en una base de datos no relacional, específicamente en MongoDB. Esta base de datos se emplea con el propósito de almacenar los datos de manera eficiente y facilitar su acceso para futuros usos.

Con el conjunto de datos obtenido en la fase anterior, se procedió a realizar una división del dataset en tres conjuntos distintos: entrenamiento, validación y prueba, tal como se evidencia en la imagen 3.3.

La distribución se definió asignando un porcentaje específico a cada uno de ellos en función de su relevancia. Se destinó el 70 % de los datos al conjunto de entrenamiento, el 20 % al conjunto de validación y el 10 % al conjunto de prueba.



Figura 3.3: Distribución de entrenamiento del modelo (Figura propia)

Selección de herramienta modelo:

El modelo empleado para llevar a cabo la clasificación de la veracidad de las noticias fue el Fine-Tuning de ChatGPT, utilizando ADA.

Este modelo de clasificación de texto se basa en la tecnología Generative Pre-trained Transformer-3, capaz de procesar y generar lenguaje natural, lo cual resulta idóneo para comprender el contexto de las noticias y evaluar sus características de veracidad. De esta manera, se logró aprovechar plenamente el potencial de las redes preentrenadas para obtener resultados más efectivos en la detección de noticias falsas.

El procedimiento llevado a cabo para el Fine-Tuning se describe a continuación: primero, se seleccionó el modelo de procesamiento de texto ADA. Luego, se ajustó la estructura del modelo preentrenado, eliminando internamente las capas que no eran pertinentes para la nueva tarea y añadiendo aquellas necesarias para la identificación específica de noticias falsas o verdaderas.

En esta nueva configuración del modelo, solo se congelaron las capas provenientes de la red original, cuyo conocimiento se deseaba preservar para el nuevo entrenamiento. Esto impidió que los pesos de las capas del modelo preentrenado cambiaran durante el proceso de entrenamiento, asegurando así la retención del conocimiento adquirido por el modelo sobre el lenguaje natural.

Entrenamiento, validación y testeo del modelo:

Durante esta etapa, el modelo se entrena con dos campos esenciales: el primero es el campo “prompt”, que contiene la información de la noticia, y el segundo es el campo “completion”, que define la etiqueta correspondiente de la noticia como falsa o verdadera. Utilizando estos datos previos, el modelo es entrenado mediante un método de optimización, una función de pérdida y una métrica de evaluación. En el transcurso del entrenamiento, únicamente se ajustan los pesos de las capas no fijas para adaptarse a la nueva tarea.

El conjunto de datos de entrenamiento se emplea para instruir al modelo de aprendizaje automático. A lo largo de este proceso, el modelo ajusta sus paráme-

tros para aprender patrones y características relevantes presentes en los datos. Este conjunto de datos de entrenamiento es lo suficientemente extenso y representativo para permitir que el modelo identifique patrones comunes en las noticias falsas, abordando aspectos como la redacción y estructura de las noticias, entre otros.

Para garantizar la precisión en la detección de noticias falsas, se seleccionó un conjunto de datos grande y diverso, asignando un 70 % de los datos al proceso de entrenamiento. Esta elección se fundamenta en la necesidad de proporcionar al modelo una cantidad suficiente de información para aprender patrones relevantes, sin comprometer la eficiencia del proceso de entrenamiento.

El proceso de entrenamiento del modelo tuvo una duración aproximada de 4 horas, llevado a cabo mediante la API de OpenAI.

El conjunto de validación se utilizó para evaluar el modelo durante el entrenamiento, ofreciendo una evaluación imparcial del ajuste del modelo en el conjunto de datos de entrenamiento mientras se ajustan los hiperparámetros del modelo.

El conjunto de pruebas se empleó para evaluar el rendimiento del modelo después del entrenamiento, proporcionando una evaluación imparcial del desempeño del modelo en datos no vistos durante el proceso de entrenamiento.

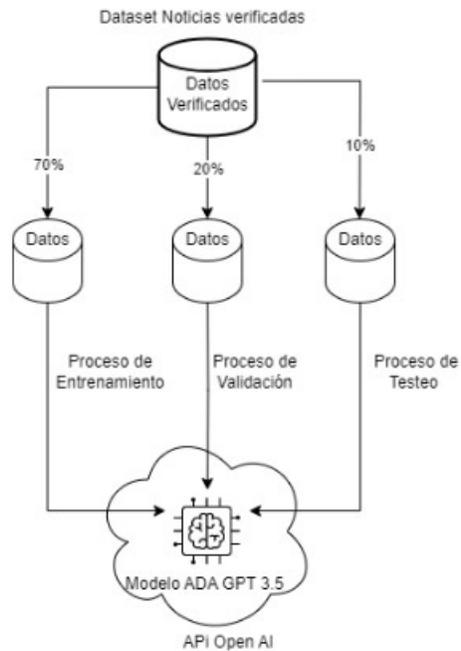


Figura 3.4: Proceso de entrenamiento del modelo (Figura propia)

La figura 3.4 ilustra el proceso de entrenamiento del modelo, incluyendo la correspondiente división de datos para cada etapa del proceso de entrenamiento.

Definición de red social twitter:

La elección de Twitter como plataforma para analizar la propagación de noticias falsas se sustenta en los siguientes fundamentos.

En primer lugar, Twitter representa un escenario dinámico y de rápido flujo de información, donde las noticias en el contexto político, se difunden de manera instantánea, permitiendo un seguimiento en tiempo real de eventos y conversaciones. Además, la extensa base de usuarios de Twitter abarca diversas geografías, culturas y sectores, brindando una amplia muestra de datos para un análisis más completo y representativo. La naturaleza concisa de los tweets también presenta un desafío único, ya que la información debe ser comprimida en un espacio limitado, lo que puede influir en la forma en que se presenta y percibe la noticia.

La viralidad inherente de la plataforma, con retweets y menciones, facilita el rastreo de la difusión de noticias y la identificación de patrones de propagación. En resumen, la elección de Twitter como objeto de estudio para analizar la propagación de noticias falsas se justifica por su inmediatez, diversidad de usuarios y la peculiaridad de su formato, ofreciendo un terreno fértil para comprender y abordar la problemática de la desinformación en línea.

Definición de cuentas:

Las particularidades y rasgos distintivos del contexto político colombiano ejercen una notable influencia en la difusión de noticias falsas. En consecuencia, la propuesta de un sistema de alerta temprana se ha ajustado y adaptado para abordar de manera específica estos elementos, focalizándose en el entorno político de Colombia.

En este sentido, se identificaron una serie de factores que orientan la selección y determinación de la relevancia de las cuentas en la red social twitter, de un individuo en el ámbito político. Los factores son los siguientes:

1. **Cargo o posición en el ámbito político:** Determinar el cargo, posición o rol que ocupa la persona o entidad en el ámbito político, incluyendo categorías como líder, candidato, funcionario, partido político, organización, medio de comunicación, influencer político, entre otros.
2. **Nivel de influencia en el ámbito político:** Evaluar el nivel de influencia que posee la persona o entidad en el ámbito político, considerando aspectos como poder de decisión, opinión, movilización, negociación, entre otros.
3. **Temas o agendas políticas promovidos o defendidos:** Identificar y analizar temas o agendas políticas que la persona o entidad promueve o defiende, discerniendo si son de interés nacional, regional, local, sectorial, u otros.
4. **Tipo de contenido en la cuenta de Twitter:** Clasificar el tipo de contenido publicado por la persona o entidad en su cuenta de Twitter, considerando si es informativo, opinativo, persuasivo, polémico, u otros.
5. **Tono o estilo en la cuenta de twitter:** Analizar el tono o estilo empleado por la persona o entidad en su cuenta de Twitter, evaluando si es formal, informal, respetuoso, irrespetuoso, constructivo, destructivo, entre otros.
6. **Tipo de interacción con seguidores y otros actores políticos:** Caracterizar el tipo de interacción que la persona o entidad mantiene con sus seguidores y otros actores políticos en su cuenta de Twitter, determinando si es activa, pasiva, receptiva, hostil, colaborativa, confrontativa, etc.
7. **Reacción generada por las publicaciones en twitter:** Evaluar la reacción que generan las publicaciones de la persona o entidad en su cuenta de Twitter, distinguiendo si es positiva, negativa, neutra, mixta, entre otras posibles respuestas.
8. **Riesgo de información en la cuenta de twitter:** Analizar el tipo de riesgo de información que la persona o entidad podría representar a través de su cuenta de Twitter, considerando la posibilidad de divulgar información sensible, falsa, manipulada, confidencial, entre otros

A partir de los anteriores criterios, se llevó a cabo una investigación con el objetivo de identificar diversos actores políticos, medios de comunicación y entidades que pudieran ostentar relevancia e influencia en la esfera política nacional a través de la red social Twitter.

Los datos seleccionados para cada cuenta incluyeron:

Tabla 3.1: Datos seleccionados para cada cuenta

Nombre	Descripción
Usuario (Screen Name)	Nombre de la cuenta del usuario
Nombre de usuario	Nombre del dueño de la cuenta
Seguidores	Cantidad de seguidores de la cuenta
Descripción	Descripción de la cuenta

En concordancia con los criterios establecidos en la tabla 3.1, se identificaron un total de 93 cuentas, las cuales engloban datos vinculados a distintos políticos destacados, periodistas, plataformas de noticias y figuras relevantes en el ámbito político nacional.

Estas cuentas fueron sometidas a un análisis detallado y se emplearon para recopilar información diversa de los tweets.

Recolección de medios no verificados:

La obtención de tweets se llevó a cabo mediante la API v2 de Twitter, la cual implicó un costo mensual de 100 dólares. Este proceso se implementó a través de la librería Tweepy de Python, con la construcción del script en un entorno de Jupyter Notebook para facilitar el manejo y procesamiento de la información. La elección de utilizar la API de Twitter se fundamenta en las actualizaciones recientes de la política de datos, que limitan la extracción de información mediante técnicas convencionales de web scraping, especialmente en el caso de noticias provenientes de fuentes verificadas.

A continuación, se presenta el diagrama 3.5 de consulta que guió este procedimiento:



Figura 3.5: Diagrama de consulta

Con los parámetros previamente establecidos para cada cuenta, se efectuó una solicitud tipo “GET” a la API de Twitter con el fin de obtener la información correspondiente de cada cuenta.

En la tabla 3.2 se presentan los datos obtenidos:

Tabla 3.2: Datos obtenidos

Campo	Descripción
Usuario	Nombre de usuario en Twitter
Identificación	Id de usuario en Twitter
Nombre	Nombre real del usuario
Ubicación	Ubicación geográfica del usuario
Descripción	Descripción proporcionada por el usuario en su perfil
Entidades	Entidades con las cuales el usuario a interactuado
Número de seguidores	Cantidad de contactos o usuarios a los que el perfil sigue.
Número de amigos	Descripción de la cuenta
Fecha de creación	Fecha y hora en que se creo la cuenta
Listas públicas	Informa si la cuenta del usuario cuenta con la verificación de Twitter.
Verificado	Descripción de la cuenta

Recolección de tweets de cuentas seleccionadas:

Con las cuentas previamente seleccionadas en la etapa anterior, se llevó a cabo la recopilación de información de los tweets más recientes publicados por cada

cuenta. Este procedimiento se ejecuta mediante el ID de usuario de cada cuenta, mediante el cual se solicitan diversos metadatos correspondientes a cada tweet. Este enfoque se adoptó para llevar a cabo un estudio en profundidad. Los campos específicos solicitados fueron los siguientes:

Tabla 3.3: Campos específicos solicitados

Campo	Descripción
Tweet ID	Identificador único de tweet
Text	Texto de tweet
ID Author	Identificador único del usuario que publicó el tweet
Publication date	Fecha de publicación del tweet
retweet count	Número de veces que ha sido retwitteado
reply count	Número de veces que ha recibido una respuesta
like count	Número de veces que ha recibido me gusta
quote count	Número de veces que ha sido citado
bookmark count	Número de veces que ha sido guardado
impression count	Número de veces que ha sido visto.

A continuación, se presentan ejemplos de datos extraídos mediante la API:

Campo	Tweet 1	Tweet 2	Tweet 3	Tweet 4
Tweet ID	1691767009233563783	1686481665332219904	1685352038043901952	1683458611757072384
Text	Si alguna de las partes...	Este jueves a las 6:00 pm...	Lamento mucho el fallecimiento...	Felicitaciones a nuestra querida...
Id autor	64839766	64839766	64839766	64839766
Date	2023-08-16	2023-08-01	2023-07-29	2023-07-24
Retweet	726	47	106	38
Reply	1783	68	44	32
Like	3553	283	688	337
Quote	299	7	9	1
Bookmark	14	3	3	1
Impression Count	379022	66982	68810	24653

Tabla 3.4: Resumen de Tweets.

Los identificadores (IDs) de los tweets están asociados al expresidente Juan Manuel Santos, abordando una diversidad de temas que abarcan desde cuestiones políticas y sociales hasta acontecimientos personales. Destaca, entre ellos, el tweet con el mayor número de interacciones, publicado el 16 de agosto de 2023,

donde el expresidente responde a acusaciones de corrupción dirigidas hacia él. Este mensaje obtuvo un notable alcance con 726 retweets, 1783 respuestas, 3553 me gusta y 299 citas.

Contrastando, el tweet con el menor número de interacciones corresponde al 24 de julio de 2023, en el cual el expresidente extiende sus felicitaciones a la Armada Nacional de Colombia por su bicentenario. Este tweet registró 38 retweets, 32 respuestas, 337 me gusta y 1 cita.

Criterios para selección de tweet a validar:

Con el propósito de llevar a cabo la validación y selección de tweets susceptibles de generar un impacto político en la población, se implementó la siguiente fórmula de clasificación. Este enfoque se fundamenta en el análisis detallado proporcionado por el estudio de referencia [54], el cual se apoya en tres modalidades de medición.

Cantidad de Visualizaciones del Tweet:

Este parámetro cuantifica de manera efectiva la frecuencia con la que los usuarios han visualizado el tweet, proporcionando así una métrica objetiva para evaluar la difusión de la información. Se calculó obteniendo la media de visualizaciones de los tweets recopilados y del usuario.

Reacciones del Tweet:

Permite cuantificar el impacto en la red midiendo la cantidad total de reacciones, que incluye respuestas y retweets.

$$\text{Cantidad de Reacciones del Tweet} = \text{replies} + \text{retweets}$$

Impacto Relativo del Tweet:

El impacto relativo del tweet establece una medida de cuantificación del efecto en relación con las interacciones pasadas de la cuenta [55]. Se mide de la siguiente manera :

$$\text{Reacciones del Tweet} = \text{replies} + \text{retweets}$$

Valor Esperado de Reacciones = promedio histórico de reacciones desde la cuenta propagadora

$$\text{Impacto Relativo} = \frac{\text{Reacciones del Tweet}}{\text{Valor Esperado de Reacciones}}$$

Este análisis se aplicó a cada uno de los tweets recopilados de las cuentas estudiadas. El valor esperado de las reacciones se obtuvo mediante el promedio de

todos los tweets recopilados.

De esta manera, se obtuvieron el valor esperado de reacciones y el impacto relativo para cada cuenta, así como el promedio de visualizaciones o impresiones de tweets. Los resultados se presentan en la tabla 3.5.

Usuario	Nombre	Valor Esperado	Cantidad de Impresiones
ABenedetti	Armando Benedetti	951.53	228717.03
IvanDuque	Iván Duque	1632.59	156229.66
petrogustavo	Gustavo Petro	1959.85	133752.46
VickyDavilaH	Vicky Dávila	2289.53	103990.14
AndresPastrana_	Andrés Pastrana A	1800.88	99361.80
QuinteroCalle	Daniel Quintero Calle	1038.99	98071.89
nalcocer_petro	Nicolas Alcocer Petro	2374.82	90979.31
sergio_fajardo	Sergio Fajardo	573.37	88882.08
GustavoBolivar	Gustavo Bolívar	3146.41	85289.99
intiasprilla	Inti Asprilla	3027.46	79437.29
JuanManSantos	Juan Manuel Santos	1102.25	79243.13

Tabla 3.5: Resumen de Valor Esperado y Cantidad de Impresiones por Usuario.

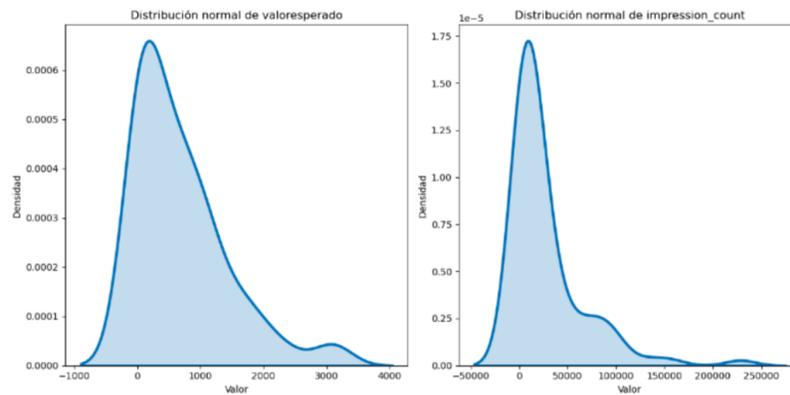


Figura 3.6: Gráfica de distribución normal (Figura propia)

En la figura 3.6, se observa que la media del valor histórico de tweets compartidos es de 552, mientras que la cantidad de visualizaciones de un tweet alcanza las 18,552 veces. Estos valores sirven como referencia para establecer un punto comparativo con los tweets de cada cuenta durante el proceso de validación.

A partir de los datos recopilados, se llevó a cabo un análisis detallado para identificar las cuentas que ejercen una mayor influencia en la red, permitiendo así examinar críticamente la información que comparten, especialmente en el ámbito político donde sus opiniones pueden tener un impacto más significativo.

Este análisis se enfoca en dos métricas clave:

El valor promedio esperado de reacciones y las visualizaciones promedio de cada tweet por cuenta. La utilización del percentil en el cálculo de estas variables proporciona una evaluación objetiva de la posición relativa de cada cuenta en el conjunto de datos, revelando cuáles de ellas poseen un peso más considerable en la red [56].

El cálculo del percentil se realiza ordenando todos los valores en el conjunto de datos de manera ascendente y determinando la posición relativa dentro del conjunto. Esta posición se obtiene al tomar el conjunto de valores que son menores al valor dado, dividirlo entre el número total de valores y multiplicarlo por 100, según se muestra en la siguiente fórmula:

$$P = \frac{R}{N} \times 100$$

Donde P es el percentil, R es el rango del valor (la posición del valor cuando los datos están ordenados ascendentemente) y N es el número total de datos u observaciones.

Los valores de percentil cercanos a 0 indican una influencia baja, mientras que los valores cercanos a 1 señalan un mayor grado de influencia.

Una vez calculado el percentil para cada cuenta en ambas variables, se sumaron los percentiles para obtener una medida combinada del “peso” de cada cuenta en ambas métricas.

Ordenando las cuentas en función de esta suma, se seleccionaron aquellas que superaban el umbral del 70 %, ya que en este percentil se encuentran las cuentas con un impacto significativo en la red, con más de 25,000 impresiones promedio por tweet. De las 93 cuentas inicialmente consideradas, se seleccionaron las 25 que cumplían con estos criterios.

A continuación, en la tabla 3.4 se presentan las primeras 5 cuentas seleccionadas.

A partir de la recopilación de tweets, se obtuvo un total de 2609 publicaciones. Esta información fue empleada para llevar a cabo un análisis detallado de la información compartida. El análisis se aplicó individualmente a cada tweet, utilizando como criterio el impacto relativo, determinado por el valor esperado del promedio

Usuario	Nombre	Percentil de valores esperados	Percentil de impresiones	Suma de percentiles
VickyDavilaH	Vicky Dávila	0.957447	0.968085	1.925532
petrogustavo	Gustavo Petro	0.946809	0.978723	1.925532
GustavoBolivar	Gustavo Bolívar	1.000000	0.914894	1.914894
nalcocer_petro	Nicolas Alcocer Petro	0.968085	0.936170	1.904255
IvanDuque	Iván Duque	0.893617	0.989362	1.882979

Tabla 3.6: Primeras 5 cuentas seleccionadas.

de reacciones en relación con el recuento real de reacciones de cada publicación. Cuando el impacto relativo superaba el valor de 1, indicaba que la publicación superaba el promedio de reacciones por tweet, por ende, era seleccionada para un análisis más exhaustivo.

Bajo este criterio, se identificaron un total de 771 tweets, como se representa a continuación:

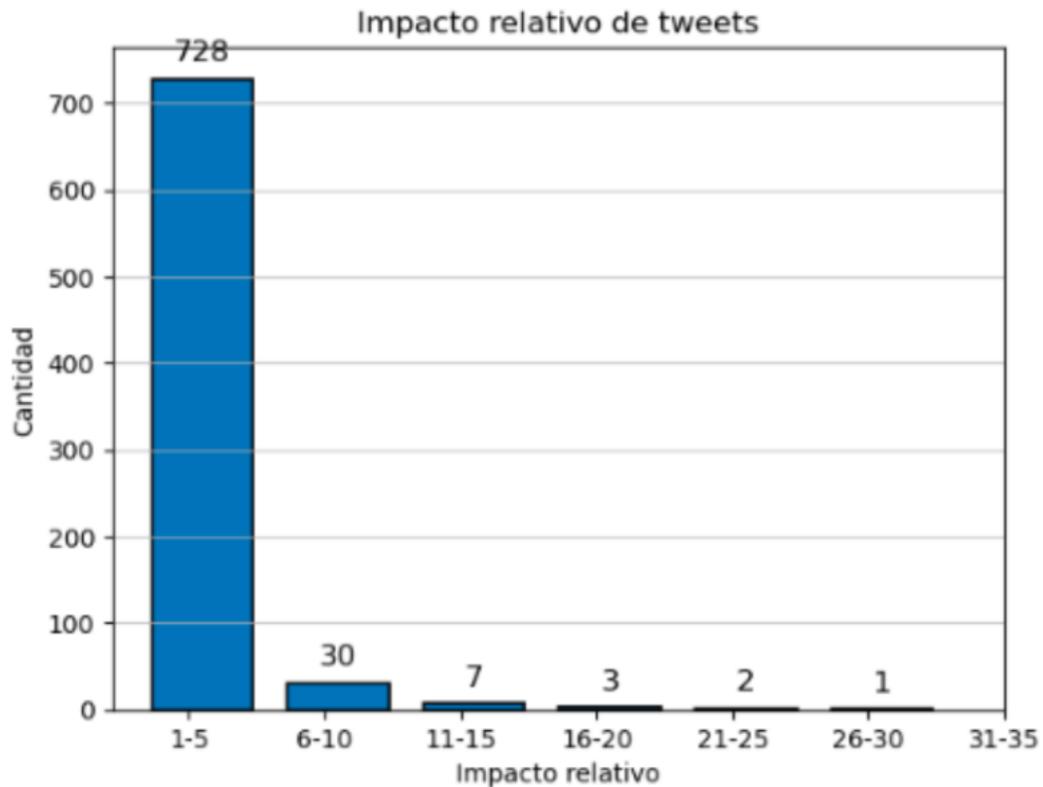


Figura 3.7: Impacto relativo de tweets (Figura propia)

La figura 3.7 revela el recuento de tweets, organizados en conjuntos de 5, en función del impacto relativo. Destaca que el grupo de 1 a 5 exhibe el mayor peso porcentual, abarcando el 94.42 % de los tweets que superan el valor esperado. El restante 5.58 % se distribuye en los demás conjuntos.

Este análisis se justifica considerando que la visibilidad de un tweet tiende a aumentar con el incremento del impacto relativo. Para propósitos de análisis, se han tomado en cuenta únicamente aquellos tweets con un impacto relativo igual o mayor a 6.

A continuación, se presenta la tabla 3.7 con los datos agrupados para una mejor comprensión de los resultados.

Intervalo de Valor	Número de Repeticiones	Porcentaje (%)
1.0 - 5.0	728	94.42
6.0 - 10.0	30	3.24
11.0 - 15.0	7	0.78
16.0 - 20.0	3	0.39
21.0 - 25.0	2	0.13
26.0 - 30.0	1	0.13
Total	771	100

Tabla 3.7: Distribución de tweets según intervalo de valor.

Los tweets seleccionados para el prototipo ascienden a 36, los cuales poseen un número de retweets igual o superior a 6 veces su valor esperado.

Dada la rápida capacidad de difusión y viralización que estos tweets pueden alcanzar en la red, se realizó el análisis utilizando un modelo previamente entrenado. En este proceso, se introdujo el texto de la noticia en la solicitud del modelo, el cual lleva a cabo un análisis interno para determinar tanto el grado de veracidad de la noticia como la probabilidad asociada.

A continuación, en la tabla 3.8, se presentan los resultados obtenidos de prototipo para 5 de los 36 tweets seleccionados.

Noticia	Veracidad	Probabilidad (%)
Fiscales revelaron cómo cayeron en la ‘trampa’ de policías de la Dijin para autorizar interceptaciones ilegales en el caso de Laura Sarabia. Tienen miedo y denuncian intimidaciones	Falsa	95
Urgente: fiscal Francisco Barbosa denuncia campaña de desprestigio contra vicefiscal Martha Mancera; señala al Gobierno Petro, “me la quieren volver mafiosa”	Falsa	97
Este es el polémico decreto del Gobierno para las Fuerzas Armadas, en medio del cese al fuego con las Farc: tienen que cuidar a los terroristas y garantizarles el paso de alimentos	Falsa	99
Solo un antisemita puede sostener que el presidente Gustavo Petro no fue antisemita al atacar a los judíos y compararlos con los nazis y el monstruo Hitler y a Gaza con el Holocausto.	Opinión	N/A
Por qué es tan difícil de entender esto: Es una obligación moral de los Estados y los seres humanos (presidentes, líderes y ciudadanos) condenar desde cualquier lugar del mundo el sangriento e infame ataque terrorista de Hamás contra Israel. La masacre de bebés por parte...	Verdadera	85

Tabla 3.8: Análisis de Veracidad de Noticias.

Mediante esta herramienta, es factible analizar diversos tipos de noticias. No obstante, cobra una relevancia crucial llevar a cabo un análisis exhaustivo de

los datos previos.

La importancia de prototipo radica en que, de acuerdo con los resultados obtenidos, se puede determinar el nivel de influencia que los tweets ejercen sobre la información compartida.

CAPÍTULO 4

FASE DE DESARROLLO DEL MODELO CONCEPTUAL



4.1. Introducción

Este capítulo aborda la implementación práctica de los conceptos teóricos y metodologías establecidas en los capítulos anteriores.

A través de un enfoque meticuloso, se explora la creación y refinamiento del prototipo, destacando los desafíos enfrentados y las soluciones innovadoras adoptadas. Se presenta un análisis de las características, la arquitectura y la funcionalidad del prototipo, evidenciando así la materialización de la propuesta conceptual en una herramienta tangible y funcional.

De este modo, la Figura 4.1, presenta el contenido abordado en el Capítulo 4:

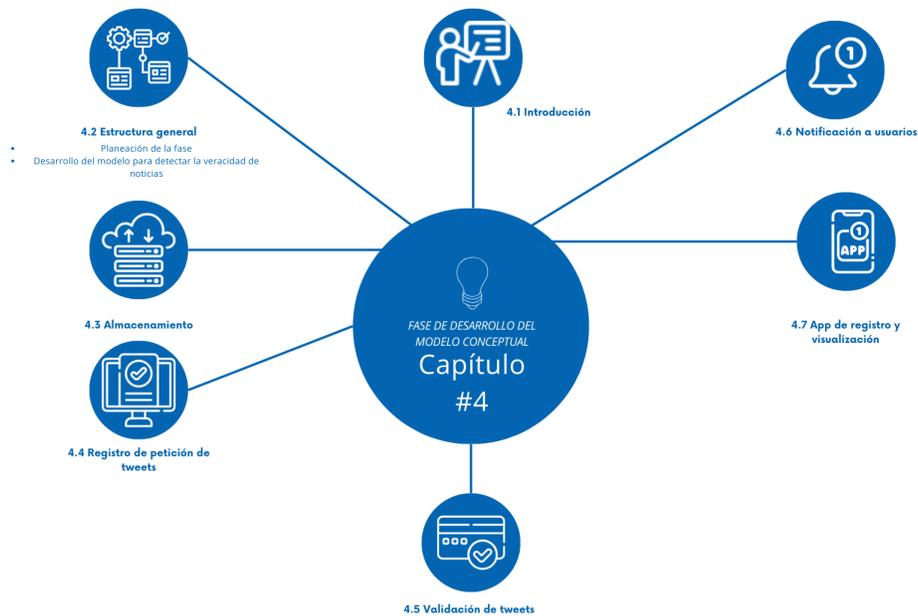


Figura 4.1: Mapa de contenido capítulo 4 (Figura propia)

4.2. Estructura general

El diseño del prototipo del sistema de alerta temprana se concibió con el propósito de ser una herramienta que mantuviera a los usuarios informados acerca de posibles noticias falsas o desinformación relacionada con temas políticos circulando en Twitter.

En la construcción, se implementó el proceso delineado en el capítulo 3, que

abordó la clasificación, selección y validación de tweets. Este proceso fue integrado en un aplicativo basado en FastAPI, junto con otros microservicios que, en conjunto, conforman todo el prototipo del sistema de alerta temprana. Estos servicios incluyen microservicios de almacenamiento, servicios de notificación y una aplicación móvil como interfaz de interacción de usuarios.

En la figura 4.2 se presenta el diagrama de componentes del sistema.

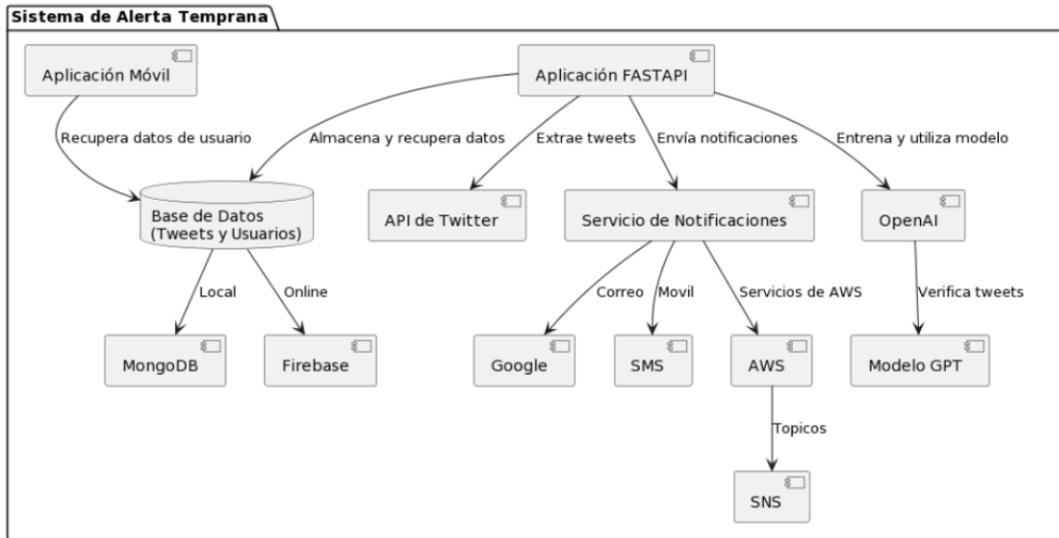


Figura 4.2: Diagrama de componentes del sistema (Figura propia)

El sistema se concibió bajo una arquitectura mixta que fusiona una aplicación web desarrollada con el framework FASTAPI y una aplicación móvil como interfaz de usuario. La aplicación web integra servicios clave, como la API de Twitter, encargada de extraer, verificar y notificar tweets clasificados y previamente seleccionados de usuarios relevantes. También incorpora un modelo de generación de lenguaje natural (GPT), previamente entrenado en OpenAI, para verificar la veracidad de los tweets mediante la comparación de contenido y estructura.

Para las notificaciones, se emplearon tres métodos: el servicio SNS de AWS para enviar notificaciones a usuarios registrados, los servicios de API de correo electrónico de Google como segundo método, y la utilización de mensajes SMS a través de la aplicación como tercer método.

El sistema de almacenamiento de información se implementó en dos gestores de bases de datos NoSQL: MongoDB, donde se almacenan todos los datos en bruto solicitados desde la aplicación, y Firebase, utilizado como backend intermedio para todas las noticias verificadas y como sistema de gestión de registro rápido

de usuarios.

Por su parte, la aplicación móvil, desarrollada en kotlin, se dividió en dos partes: una para la administración, control de usuarios y noticias, y otra que muestra las noticias verificadas que cumplen con los parámetros para notificar a los usuarios.

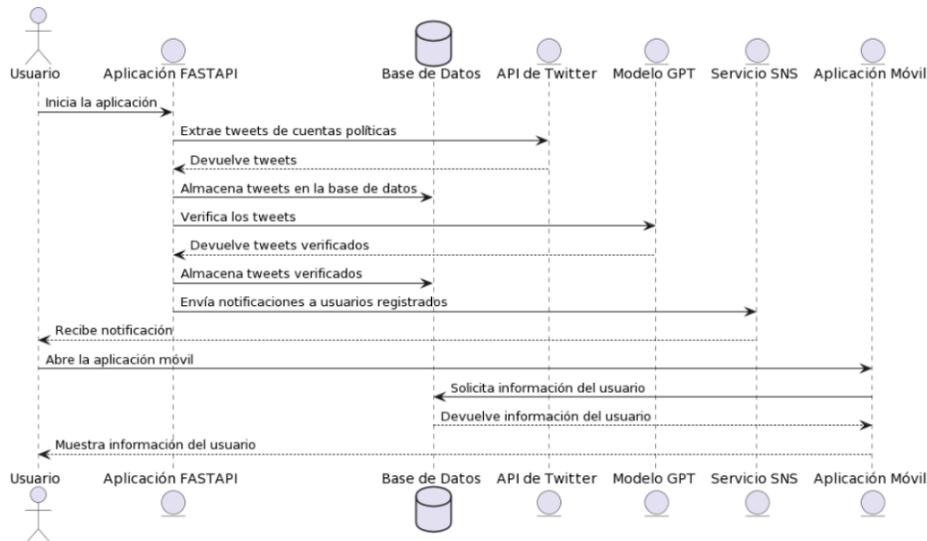


Figura 4.3: Diagrama de secuencia (Figura propia)

El diagrama de secuencia representado en la figura 4.3 detalla el flujo operativo desde la extracción inicial de los tweets hasta el procedimiento final de notificación a los usuarios.

A continuación, se presenta una descripción paso a paso de este proceso:

- El administrador inició la aplicación FASTAPI, encargada de extraer tweets de la API de Twitter y almacenar la información en una base de datos en MongoDB.
- Internamente, la aplicación llevó a cabo la validación de todos los tweets utilizando los criterios de impacto relativo descritos en el capítulo 3.
- Los tweets que superaron el proceso de filtrado y selección fueron sometidos a verificación mediante el modelo GPT, previamente entrenado.
- Posteriormente, se procedió al almacenamiento de los tweets verificados en una base de datos Firebase para facilitar la gestión y recuperación de datos.

- Se llevó a cabo un proceso diario de verificación de nuevas noticias, identificando aquellas que no habían sido notificadas o usuarios pendientes de suscripción.
- La aplicación ejecutó un proceso de envío de notificaciones a los usuarios a través de un servicio de notificaciones, que a su vez utilizó SNS para notificar a usuarios suscritos mediante correo o SMS. Simultáneamente, se emplearon los servicios de Gmail de Google y SMS para enviar notificaciones a usuarios individuales.
- Los usuarios finales recibieron las notificaciones.

4.3. Almacenamiento

Para el almacenamiento de datos, se optó por la combinación de dos sistemas robustos: MongoDB y Firebase. La elección de MongoDB como sistema de almacenamiento local se fundamentó en su capacidad para gestionar datos no estructurados provenientes de la API de Twitter, sin filtrar, además de cumplir con la función crucial de servir como respaldo integral. Por otro lado, Firebase fue seleccionado como el sistema idóneo para almacenar las noticias filtradas y operar como backend intermedio entre la aplicación FastAPI y la aplicación móvil.

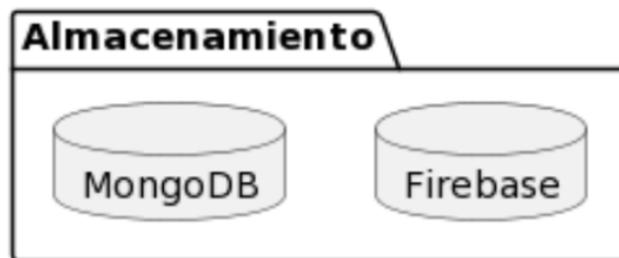


Figura 4.4: Opciones de almacenamiento (Figura propia)

MongoDB y Firebase fueron elegidos estratégicamente debido a sus características únicas y complementarias. MongoDB, como base de datos NoSQL, se destacó por su flexibilidad y capacidad para manejar grandes volúmenes de datos estructurados y no estructurados, siendo especialmente adecuado para almacenar registros sin procesar de usuarios y tweets de la API de Twitter. Por su parte, Firebase, como base de datos en tiempo real, permitió un enfoque dinámico para almacenar datos filtrados, listos para notificar a los usuarios. Además, su función como backend intermedio facilitó la sincronización de datos entre la aplicación FastAPI y la aplicación móvil.

En términos de implementación, MongoDB se desplegó principalmente en la aplicación de escritorio para el almacenamiento local de datos. Los datos relacionados con temas políticos extraídos de Twitter se procesaron y almacenaron en MongoDB para un análisis posterior. En el marco de MongoDB, se definieron esquemas de datos distintos para organizar la información y analizar patrones y características. Estos esquemas incluyeron usuarios de tweets seleccionados, permitiendo recopilar información detallada de cada cuenta de Twitter a estudiar, así como esquemas para cada tweet que recopilan los metadatos relevantes. Además, se diseñaron esquemas para usuarios del sistema, noticias verificadas y notificaciones, este último manteniendo un registro de noticias de alto impacto y verificadas, listas para ser enviadas a los usuarios.

A continuación, se presenta el diagrama de clases implementado en MongoDB para visualizar la estructura detallada del almacenamiento.

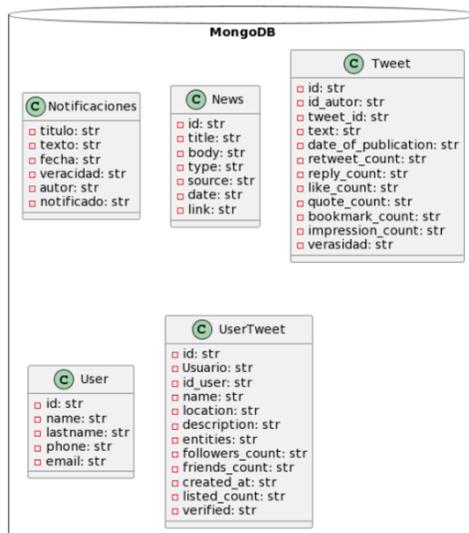


Figura 4.5: Diagrama de almacenamiento en MongoDB (Figura propia)

Firestore, por su parte, resguarda los datos de noticias que han pasado por el proceso de filtrado y están preparadas para notificarse a los usuarios. Dentro de esta información, se incluyen los esquemas que definen los datos de los usuarios del sistema, las noticias de alto impacto, donde se almacenan los tweets previamente filtrados para facilitar el proceso de notificación, y el esquema de notificaciones. Es relevante destacar que toda esta información se comparte con la base de datos de MongoDB, estableciendo una conexión efectiva entre ambos sistemas. Las operaciones ejecutadas en las tablas son sincronizadas con el respaldo local en MongoDB, proporcionando una capa adicional de seguridad. Estos datos, ya procesados y listos para su visualización, son accesibles tanto para la aplicación de escritorio en FastAPI como para la aplicación móvil desarrollada en Kotlin. A continuación, se presenta una visualización de los esquemas de datos almacenados en Firestore.

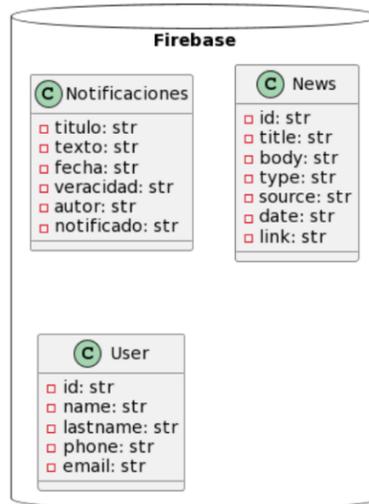


Figura 4.6: Diagrama de almacenamiento en Firestore (Figura propia)

En resumen, la combinación de MongoDB y Firebase permitió al prototipo del sistema de alerta temprana manejar eficientemente grandes volúmenes de datos, proporcionar actualizaciones en tiempo real y mantener la sincronización entre diferentes plataformas. con el fin de manejar de forma dinámica la información y poder trabajar de manera más ágil el sistema.

4.4. Registro de petición de tweets

Twitter es una red social que permite a los usuarios compartir texto, imágenes y vídeos. La API proporciona una forma de acceder a los datos, incluidos los tweets. Este procedimiento describe cómo obtenerlos utilizando la API de Twitter y la librería Tweepy.

1. Acceso a la API de Twitter Academic:

- Se solicitó acceso a la API de Twitter Academic para la investigación.
- La solicitud fue autorizada por un corto periodo de tiempo tras evaluar la finalidad académica de la investigación.

2. Acceso a la API de Twitter:

- Se utilizó la versión paga, de acceso nivel básico, la cual tiene un costo de 100 USD/mes por cada 10K tweets.
- El pago de la mensualidad se realizó para habilitar el acceso al contenido de cuentas. De otra forma la extracción no hubiese sido posible, pues existe un bloqueo a través de librerías web scrapping.

3. Utilización de la librería Tweepy:

- Tweepy es una librería de Python en la cual ya se tienen implementadas los diferentes endpoints para el acceso a la información de los usuarios.

4. Selección de los endpoints:

- Para la investigación se emplearon dos endpoints: **get user** y **get users tweets**
 - **get user:** Facilita la obtención de información del usuario a través del nombre de la cuenta o screen name. Mediante este endpoint, se accede a datos relevantes para la investigación, como el número de seguidores, la fecha de creación, información sobre la cuenta y su estado, entre otros elementos clave. Este punto de acceso es crucial, ya que posibilita la obtención del ID de la cuenta, requisito indispensable para acceder a otros datos, como el historial de tweets, entre otros.

- **get users tweets:** A partir del identificador de usuario, es posible acceder a diversos tipos de información de la cuenta. En este contexto, se facilita la obtención de un historial de los tweets más recientes, conocido como el "timeline de tweets". La solicitud incluye información sobre la consulta deseada, solicitando metadatos que resultaron fundamentales para llevar a cabo el análisis de la influencia tanto de cada cuenta como de cada tweet en la red.

4.5. Validación de tweets(Fine tuning)

La verificación de la información se efectuó mediante la implementación del modelo entrenado.

El proceso implicó la utilización del modelo de finetuning, consumido a través de la API de OpenAI. En este contexto, se suministran las credenciales de acceso correspondientes.

Posteriormente, se integra este servicio dentro de FastAPI para llevar a cabo el proceso de verificación de las noticias filtradas en Twitter.

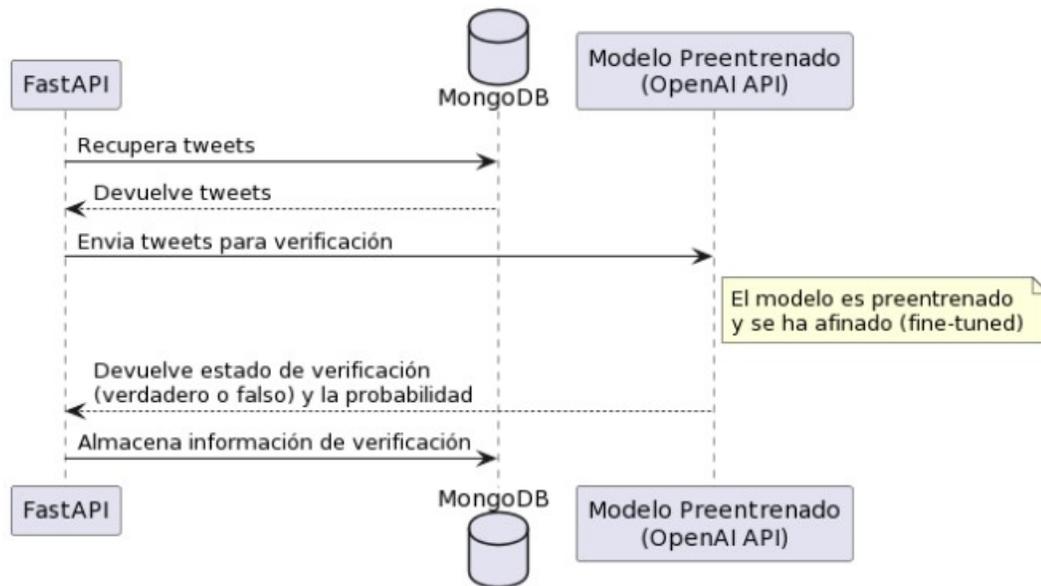


Figura 4.7: Proceso de validación de tweets (Figura propia)

En la figura 4.7 se ilustra el proceso de validación de tweets. Este inicia con la solicitud de FastAPI a MongoDB para obtener los tweets almacenados en la tabla de Notificaciones. Una vez recuperados, FastAPI los remite a la API de OpenAI para su validación. Esta API utiliza un modelo preentrenado específicamente ajustado

para la tarea en cuestión.

Procesa los tweets, verifica la información y retorna un estado de verificación que indica la veracidad o falsedad de la información, acompañado de una probabilidad asociada. La información de verificación se almacena posteriormente en MongoDB a través de FastAPI para consultas futuras. Es importante destacar que este es un esquema simplificado y que en un entorno real se gestionarían posibles errores, además de que algunas de estas operaciones se llevarían a cabo de forma asíncrona.

La solicitud a la API se realiza apuntando al modelo preentrenado, ajustado dentro del parámetro "response" en el modelo. Durante esta petición, se incluye el prompt de la consulta. Internamente, se ajustan los parámetros del modelo para garantizar un retorno preciso de caracteres, reduciendo así el consumo y minimizando la posibilidad de alucinaciones del modelo.

```
@router.post("/verificar_noticia_openai/{noticia}", tags=["Tweets"])
async def verificar_noticia(noticia: str):

    response = openai.ChatCompletion.create(
        model="detection-truev1",
        temperature=1,
        max_tokens=10,
        top_p=1,
        frequency_penalty=0,
        presence_penalty=0
    )
    return {"respuesta": response.choices[0].message['content']}
```

Figura 4.8: Script del proceso de verificación (Figura propia)

4.6. Notificación a los usuarios:

Se implementó un sistema de notificación que habilitó a los usuarios para recibir alertas mediante diversos canales, como correos electrónicos o mensajes de texto (SMS) en la aplicación móvil. La implementación de este sistema se llevó a cabo utilizando el servicio SNS de AWS, así como Gmail y la mensajería de texto convencional.

En la Figura 4.9, se presenta el diagrama de flujo correspondiente al sistema de notificaciones, haciendo uso de SNS y AWS.

En este escenario específico, se utilizó FastAPI, el cual, con las credenciales adecuadas, se conectó a través de una API Gateway con SNS. Este último se encargó de la transmisión de tópicos tanto por correo electrónico como por SMS.

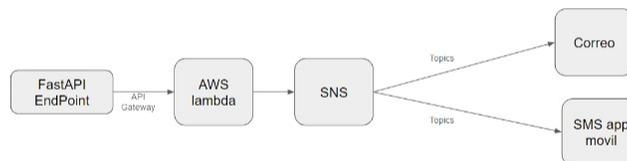


Figura 4.9: Sistema de notificaciones mediante SNS y AWS (Figura propia)

En la Figura 4.10, se presenta el diagrama de flujo correspondiente al sistema de notificaciones con Gmail.

En esta instancia, el sistema de notificación se conectó con la API de Gmail, la cual gestionó el envío de las notificaciones mediante correos electrónicos.

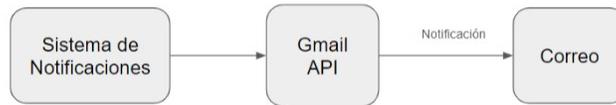


Figura 4.10: Sistema de notificaciones mediante GMAIL (Figura propia)

En la Figura 4.11, se exhibe el diagrama de flujo correspondiente al sistema de notificaciones mediante mensajes de texto.

En este escenario, el sistema de notificaciones gestionó el envío de mensajes a los usuarios a través de la aplicación móvil, utilizando el EndPoint designado para el envío de mensajes.



Figura 4.11: Sistema de notificaciones mediante SMS (Figura propia)

Los servicios elaborados se incorporaron a la aplicación, permitiendo su utilización de manera continua dentro del prototipo del sistema de alerta temprana. Esta integración se realizó a través del empleo del framework FastAPI de Python, que facilita de manera eficiente la implementación de diversos servicios esenciales para la investigación.

En el sistema, se desplegaron servicios específicos para el almacenamiento, la extracción de tweets, la verificación de la veracidad y el envío de notificaciones.

La Figura 4.12 exhibe el diagrama de clases de la aplicación desarrollada con FastAPI.

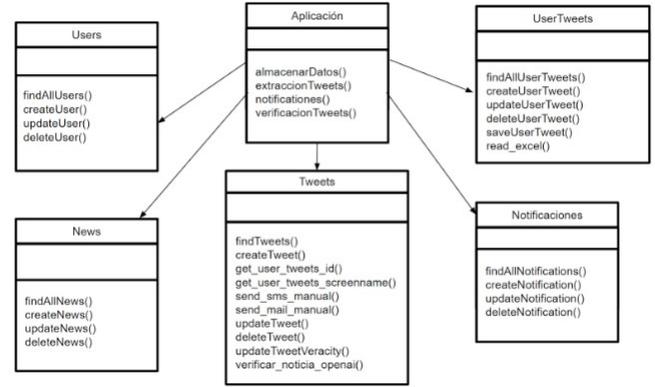


Figura 4.12: Diagrama de clases de servicios integrados en la aplicación con FastAPI.

En la figura 4.13 se muestran las etiquetas implementadas en FastAPI.



Figura 4.13: Etiquetas implementadas en FastAPI (Figura propia)

La etiqueta “Users”, de la figura 4.14, desempeña la tarea de administrar la información asociada a cada usuario. La gestión de registros y la edición de usuarios se lleva a cabo utilizando los datos proporcionados por aquellos que completaron el registro para recibir notificaciones a través de SMS y correo electrónico. Se implementaron cuatro EndPoints fundamentales para esta funcionalidad, posibilitando la ejecución de operaciones CRUD sobre los datos.

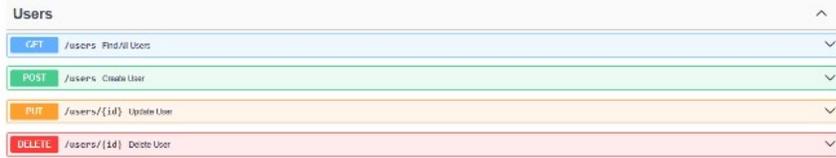


Figura 4.14: Endpoint de la etiqueta "users"(Figura propia)

La etiqueta "UserTweets", de la figura 4.15, fue implementada para proporcionar un conjunto de funcionalidades destinadas a la obtención de diversos usuarios de la plataforma Twitter a través de sus nombres de usuario, mediante una solicitud GET a la API. Además, se incorporaron métodos que permiten realizar procesos de edición o agregación de manera manual, en caso de ser necesario:

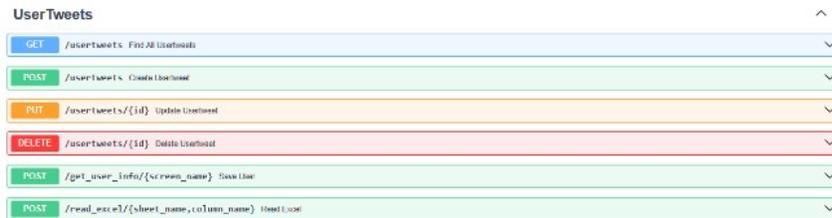


Figura 4.15: EndPoints de la etiqueta "UserTweets" (Figura propia)

Dentro de la etiqueta denominada "UserTweets", se siguió un proceso específico. Inicialmente, se efectuó una solicitud a la API de Twitter utilizando el nombre de usuario como parámetro. La API, al procesar la solicitud internamente, devolvió la respuesta solicitada en conformidad con el esquema previamente establecido. Esta petición se procesó y se almacenó como un objeto en la base de datos de MongoDB, siguiendo la estructura predeterminada. Posteriormente, con estos datos, se llevaron a cabo diversas consultas conforme a las operaciones necesarias. Además, el endpoint permitió la ejecución de operaciones manuales para agregar ciertos usuarios al sistema.

La etiqueta "Tweets", se representa la figura 4.16 cuyas operaciones permiten obtener distintos tweets de las listas de usuarios mencionadas. Se realizaban peticiones GET utilizando el identificador del usuario o mediante el screen name. Estas solicitudes permitieron obtener un historial de tweets publicados por un usuario en relación con un periodo específico. Se implementaron también funciones CRUD que facilitaron la actualización o eliminación de información según fuera necesario. Dentro de este conjunto de funcionalidades se incluyeron características adicionales como el filtrado de tweets y la verificación de su veracidad,

junto con las funciones de notificación a los usuarios.

Tweets	
GET	/tweets Find All Tweets
POST	/tweets Create Tweet
GET	/get_user_tweets_id/{id},{numero resultados} Get Tweets Id User
GET	/get_user_tweets_screenname/{screen_name,numero resultados} Get Tweets User Screenname
POST	/send_sms_manual/{numero} Create Tweet
POST	/send_facebook_manual/{numero} Create Tweet
PUT	/tweets/{id} Update Tweet
DELETE	/tweets/{id} Delete Tweet
PUT	/tweets/{id}/veracidad Update Tweet Veracity
POST	/verificar_noticia_openai/{noticia} Verificar Noticia

Figura 4.16: EndPoints de la etiqueta "Tweets" (Figura propia)

News	
GET	/news Find All News
POST	/news Create News
PUT	/news/{id} Update News
DELETE	/news/{id} Delete News

Figura 4.17: EndPoints de la etiqueta "News" (Figura propia)

Notificaciones	
GET	/Notificaciones Find All Notificaciones
POST	/Notificaciones Create Notificaciones
PUT	/Actualizar_Notificaciones/{id} Update User
DELETE	/Delete_Notificaciones/{id} Delete User

Figura 4.18: EndPoints de la etiqueta "Notificaciones" (Figura propia)

4.7. App de registro y visualización

Se desarrolló una aplicación sencilla destinada a la visualización de información, la cual brinda a los usuarios la posibilidad de acceder a noticias verificadas mediante un sistema de notificaciones, mientras exhibe el historial completo de los tweets sometidos a verificación. La aplicación se encuentra perfectamente integrada con la base de datos de Firebase.

A continuación, en las figuras 4.19 y 4.20 se presentan las vistas de la aplicación móvil para una mejor comprensión.



Figura 4.19: Interfaz del home de la app (Figura propia)



Figura 4.20: Interfaz de interacción de App (Figura propia)

La aplicación posibilita a los usuarios la visualización de todos los tweets sometidos a verificación, organizando esta información según categorías de interés, ya sea regional, nacional o vinculada a personalidades políticas. Una vez seleccionada la categoría de noticias, se pueden examinar detalladamente las noticias pertenecientes a esa clasificación.

Al navegar dentro de la aplicación, se despliega información relevante como el título, la fecha, la fuente o autor, la descripción y el estado de veracidad de cada noticia.

Además, el sistema de visualización cuenta con una aplicación de administración. Esta herramienta no solo permite la introducción de tweets verificados, sino que también sirve como medio de registro para nuevos usuarios. Esta funcionalidad fue implementada con el propósito de registrar la información necesaria para llevar a cabo el proceso de pruebas.

Las figuras 4.21 y 4.22, muestran las interfaces para registro de información y opciones de la aplicación, respectivamente.

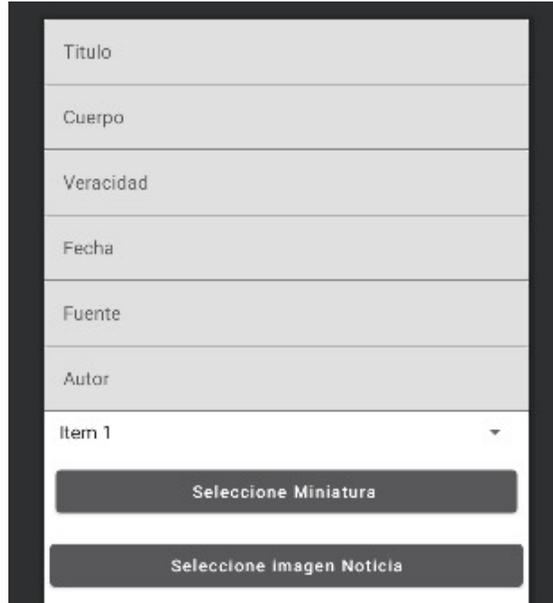


Figura 4.21: Interfaz de registro de información (Figura propia)



Figura 4.22: Interfaz de opciones disponibles en la app (Figura propia)

CAPÍTULO 5

FASE DE VALIDACIÓN DEL MODELO CONCEPTUAL



5.1. Introducción

Este capítulo describe el proceso para cumplir el objetivo 3 de la investigación, abordando la ejecución de la prueba de validación del prototipo construido.

El propósito de esta evaluación es verificar la efectividad del mismo en términos de prevención de desinformación y la diseminación de noticias falsas.

En este contexto, se proporciona una descripción de la estructura de la prueba, se documenta su ejecución, se presentan los resultados, se realiza un análisis y se identifican oportunidades de mejora.

De este modo, la figura 5.1, presenta el contenido abordado en el Capítulo 5:



Figura 5.1: Mapa de contenido capítulo 5 (Figura propia)

5.2. Prueba de validación del prototipo

La validación del prototipo se logró a través de un proceso estructurado en tres fases, las cuales se detallan en la figura: 5.2

Estructura de la prueba de validación

Sistema de alerta temprana para monitorear la propagación de noticias falsas: Caso de estudio contexto político colombiano.



Figura 5.2: Estructura de la prueba de validación

5.3. Ejecución de la prueba de validación

Para llevar a cabo la prueba, se optó por utilizar la misma cohorte de individuos que participaron en la fase inicial. La selección de usuarios se fundamentó en los siguientes criterios:

1. **Experiencia con Twitter o exposición a información de la red:** Se eligieron individuos que hubieran utilizado Twitter con anterioridad o que estuvieran expuestos a algún tipo de información proveniente de esta red.
2. **Edad entre 18 y 35 años:** Este rango etario abarca el 55,6 % de los usuarios de Twitter en Colombia. Además, coincide con el porcentaje de usuarios de las plataformas de Meta en el país [57].
3. **Edad mínima de 18 años:** Se requirió que los participantes tuvieran al menos 18 años, ya que, a esta edad, los ciudadanos colombianos tienen la capacidad de ejercer la participación política [58],[59].
4. **Potencial exposición a mensajes políticos en Twitter durante períodos electorales:** Se consideró la posible exposición de los participantes a mensajes políticos en Twitter durante los periodos electorales.

Estos criterios se justifican por las siguientes razones:

- **Usuarios activos y exposición a mensajes políticos:** La elección del rango de edad entre 18 y 35 años se basa en que estos usuarios son los más activos en Twitter y, asimismo, son quienes más se exponen a mensajes políticos en esta plataforma [59].

- **Participación política:** Ciudadanos colombianos de 18 años o más pueden ejercer la participación política, lo que los hace potencialmente influenciados por los mensajes políticos en Twitter [59].

Por ende, los usuarios de entre 18 y 35 años se consideran ideales para realizar una encuesta sobre tendencias políticas en Twitter, ya que representan el futuro de la política en Colombia. Cabe destacar que en este estudio no se tuvo en cuenta el género de los usuarios.

Ahora bien, el desarrollo de la prueba se llevó a cabo de la siguiente manera:

1. Se proporcionó una explicación básica del prototipo del sistema de alerta temprana.
2. Se presentó la aplicación móvil mediante la cual se desarrolló la prueba.
3. Se solicitó al usuario que, de la lista de 10 noticias, seleccionara una, la evaluara y expresara su opinión en relación con la veracidad de la información presentada.

Un ejemplo de la presentación de las noticias se evidencia en la figura 5.3



Figura 5.3: Ejemplo de noticias con etiquetas

4. El usuario fue notificado, ya sea a través de un mensaje de texto (SMS) o de un correo electrónico, y se le solicitó evaluar el contenido del mensaje, reflexionando sobre su relación con la noticia previamente evaluada. Un ejemplo de cómo se presentan las notificaciones se evidencia en la figura 5.4

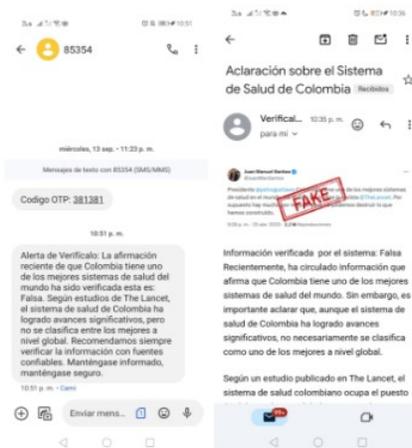


Figura 5.4: Ejemplo de notificaciones enviadas

5. Se puso a disposición del usuario el formulario de preguntas, el cual abarcó algunas preguntas del formulario CSUQ diseñado para evaluar la usabilidad de sistemas informáticos [60]. El propósito fue analizar la percepción del usuario al interactuar con el prototipo del sistema.

5.4. Análisis de resultados

La evaluación de usabilidad del prototipo de sistema de alerta temprana para mitigar la desinformación en Twitter, mediante el cuestionario CSUQ [60], se realizó en dos etapas.

La primera etapa evaluó la aplicación de interfaz visual, donde los usuarios miran la información. La segunda etapa evaluó el rendimiento del sistema de alerta temprana. Estas etapas permitieron evaluar características esenciales que permitieron comprender cómo los usuarios interactuaron con el sistema.

Según [60], se abordaron tres factores localizados: Calidad del sistema, calidad de la información y calidad de la interfaz. Las preguntas del test y su distribución se muestran en el apéndice C.

A continuación en la figura 5.5 se muestran los resultados de la evaluación de la aplicación.

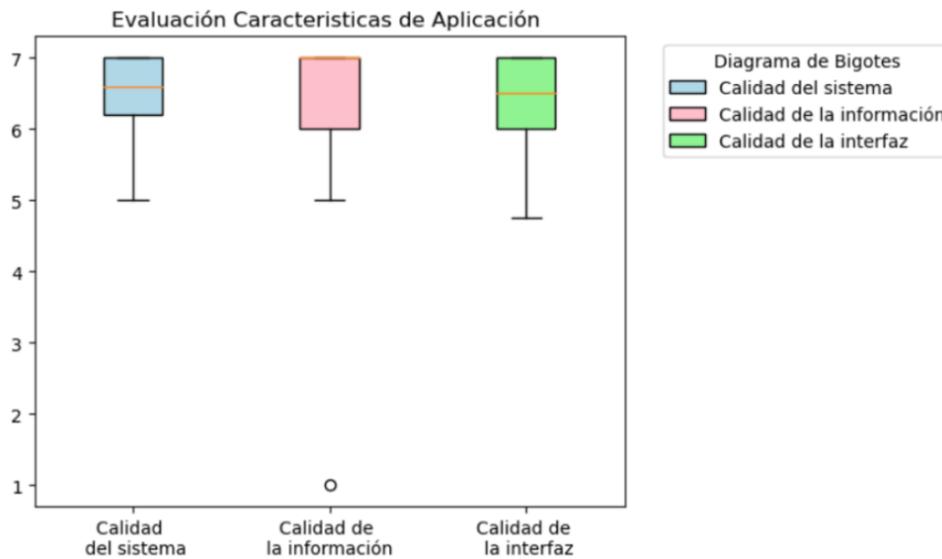


Figura 5.5: Gráfico de bigotes Características de aplicativo de visualización (Figura propia)

En la figura 5.5, los valores próximos a 1 indican insatisfacción entre los usuarios, mientras que aquellos cercanos a 7 denotan satisfacción con el uso del sistema. En relación a la calidad de la interfaz, al analizar la mediana de los datos establecida en $Q2=6.5$, con una desviación estándar de 0.61 y una media de 6.43, se observa que el valor mínimo comienza en 4.75, indicando una aceptación elevada por parte de los usuarios. Además, la percepción general se sitúa cerca de los valores medios, sugiriendo ausencia de insatisfacción en el contenido de la interfaz.

Los datos reflejan que la mayoría de los usuarios están satisfechos con el contenido de la interfaz, dado que tanto la media (6.43) como la mediana (6.5) se encuentran próximas al valor máximo de satisfacción (7). Aunque la desviación estándar (0.61) señala alguna variabilidad en las respuestas, el hecho de que el valor mínimo sea 4.75 indica que incluso el usuario menos satisfecho calificó la interfaz por encima del punto medio de la escala, sugiriendo una recepción general positiva por parte de los usuarios.

En relación a la calidad de la información, se examina $Q2=7.0$, con una media de 6.3 y desviación estándar de 0.99. Aunque esta desviación indica cierta variabilidad en las respuestas, se mantiene por encima del mínimo de 6, indicando una aceptación positiva de los datos y una satisfacción general de los usuarios con la calidad de la información proporcionada por la aplicación. A pesar de ello, se identifica un valor atípico en 1, señalando al menos un usuario insatisfecho, atribuible a factores como la diversidad de edades y la escasa familiaridad con la forma de representación de la información o el limitado uso de aplicaciones

móviles para la lectura de noticias.

Respecto a la calidad del sistema, los datos estadísticos muestran una mediana (Q2) de 6.6, una media de 6.54 y una desviación estándar de 0.52. Estos indicadores sugieren una aceptación elevada de la calidad del sistema, ya que tanto la media como la mediana se sitúan cerca del valor máximo de satisfacción (7). Aunque la desviación estándar (0.52) apunta a cierta variabilidad en las respuestas, la densidad de datos entre los cuartiles Q1 y Q2 es mayor que en las características anteriores, revelando una percepción mejorada de este parámetro. Con un valor mínimo de 5, se indica que incluso los usuarios con ciertos grados de insatisfacción perciben adecuadamente el sistema.

En términos generales, la aplicación fue bien recibida por los usuarios en lo que respecta a la calidad de la interfaz, la información y el sistema, con una amplia aceptación. Sin embargo, se identifican áreas de mejora en cuanto a la información, y se sugiere realizar una investigación más profunda con los usuarios para perfeccionar otros aspectos.

A continuación en la figura 5.6 se presentan los resultados de la evaluación del sistema de alerta temprana.

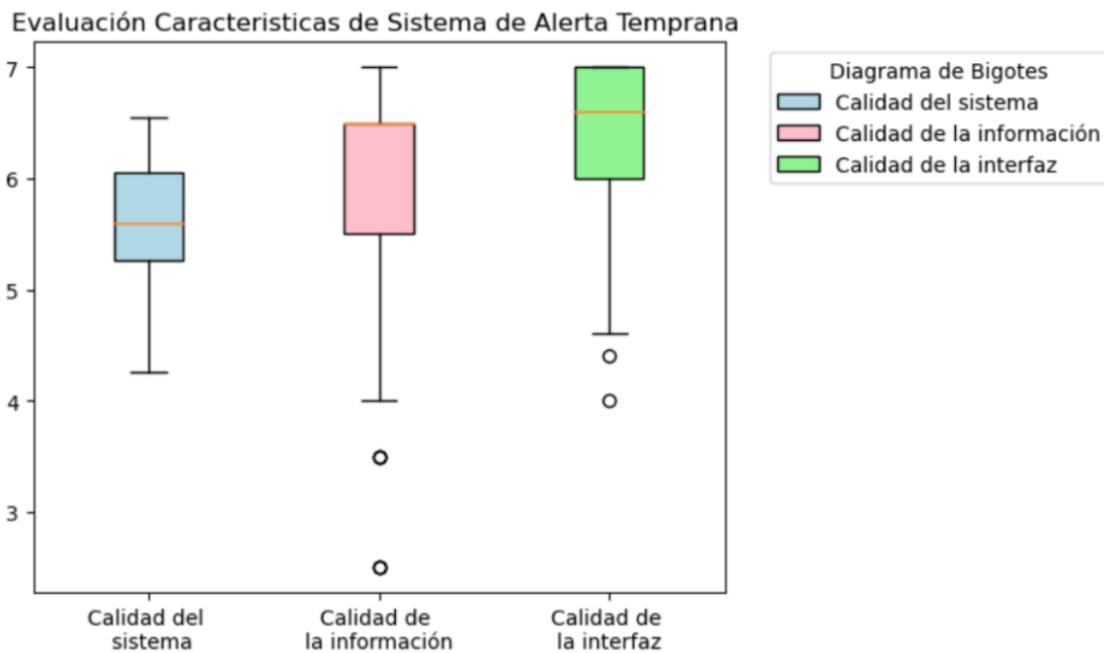


Figura 5.6: Gráfico de bigotes Evaluación de características del sistema de alerta temprana(Figura propia)

La figura 5.6, que representa la calidad de la interfaz, revela los siguientes datos: $Q2=6.6$, $media=6.6$, y valor mínimo de 4.54. Estos indicadores sugieren que los datos están distribuidos cercanos a la media, con $Q1=6$ y $Q3=7$, evidenciando una aceptación positiva de la interfaz. Sin embargo, algunos usuarios expresan insatisfacción con ciertas características visuales, como se refleja en el valor mínimo y los valores atípicos de 4 y 4.25.

En cuanto a la calidad de la información, el análisis de datos revela que $Q2$ y $Q3$ son iguales (6.5), mientras que $Q1$ es 5.5. Este patrón indica que al menos el 50% de los datos se sitúa en este rango. Además, el 25% de los datos oscila entre 6.5 y el valor máximo de 7, mientras que el 25% restante se encuentra entre 6.5 y el valor mínimo de 4. Se identificaron valores atípicos en 2.5 y 3.6. Con una desviación estándar de 1.27 y una media de 5.96, se evidencia una buena calidad de la información, bien aceptada por los usuarios, generando una percepción positiva. No obstante, existe una opinión moderada que abarca el 25% de la población encuestada entre el valor mínimo y $Q1$, indicando un nivel moderado de conformidad con la información proporcionada.

En relación a la calidad del sistema, la evaluación general muestra una disminución en comparación con los otros parámetros. La concentración entre el valor mínimo y máximo se sitúa entre 4.26 y 6.54, con una mediana de 5.59. El 50% de los datos se encuentra entre 5.26 y 6.04, indicando una percepción positiva del sistema, superando la escala media. Sin embargo, se hace necesario realizar una investigación para identificar los aspectos que requieren mejoras, con el objetivo de elevar la calidad general del sistema.

En términos generales, el sistema de alerta temprana es bien recibido por los usuarios, con puntuaciones altas en sus características, reflejando una amplia aceptación. No obstante, se destacan algunas áreas que demandan mejoras para optimizar la experiencia del usuario en cada aspecto del sistema.

A continuación, se presenta un análisis detallado de los diversos resultados de las preguntas formuladas en la encuesta, con el propósito de evaluar individualmente el grado de aceptación.

En el análisis a través del cuestionario CSUQ, se revela que el 57,9% de los participantes expresó estar "totalmente de acuerdo con la afirmación "Estoy satisfecho con el sistema", mientras que el 33,3% respondió "de acuerdo", destacando así el nivel general de satisfacción con el sistema.

Al abordar aspectos específicos relacionados con la usabilidad de la aplicación, los usuarios emitieron opiniones positivas, indicando que el sistema es:

1. Simple de usar (66,7% de respuestas "totalmente de acuerdo").
2. Cómodo de usar (64,9% de respuestas "totalmente de acuerdo").

3. Fácil de aprender a usar (63,2 % de respuestas "totalmente de acuerdo").
4. Rápido de entender su finalidad (57,9 % de respuestas "totalmente de acuerdo").
5. Capaz de proporcionar las ayudas necesarias para completar tareas (52,6 % de respuestas "totalmente de acuerdo").
6. Altamente placentero en su uso (63,2 % de respuestas "totalmente de acuerdo").
7. Agradable de usar (63,2 % de respuestas "totalmente de acuerdo").

No obstante, se identificó que el 12,3% de los participantes respondió "en desacuerdo" con la afirmación "El sistema cuenta con todas las herramientas que esperaba que tuviera".

En términos generales, los resultados de la evaluación de usabilidad indican que el prototipo es fácil de usar y satisface las expectativas de la mayoría de los usuarios.

Por otro lado, al analizar los resultados de la evaluación de efectividad, se destaca que el sistema tiene un impacto positivo en la capacidad de los usuarios para identificar noticias desinformativas, motivándolos a buscar más información y percibir la información de manera más crítica. Sin embargo, se observa que hay un sector significativo de la población que percibe el sistema como igualmente efectivo que otros métodos para evitar la desinformación, como se ilustra en la figura 5.7.

¿Cómo evalúa la efectividad del sistema de alerta temprana en comparación con otros métodos de combatir la desinformación?
57 respuestas

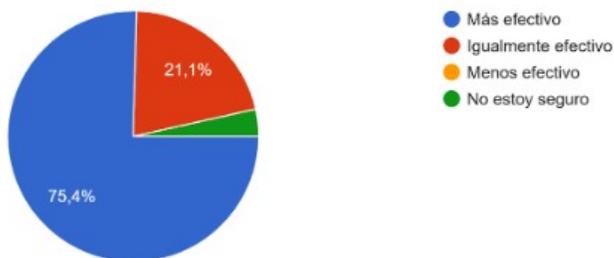


Figura 5.7: Gráfica de resultados para la pregunta "¿Cómo evalúa la efectividad del sistema de alerta temprana en comparación con otros métodos de combatir la desinformación?"(Figura propia)

Específicamente, un 40,4 % de los participantes señaló una mejora en su habilidad para identificar noticias desinformativas tras utilizar el sistema. Asimismo, un 57,9 % indicó que el sistema los incentivó a buscar información adicional sobre las noticias que recibieron. Por último, un 78,9 % expresó que el sistema influyó en su percepción y comprensión de la información.

En cuanto a la velocidad de detección de la desinformación, el 54,4 % de los participantes consideró que el sistema es rápido, mientras que el 29,8 % lo calificó como muy rápido.

Los participantes también coincidieron en que la personalización de las notificaciones y la participación activa de los usuarios son elementos clave para potenciar la efectividad del sistema. Específicamente, un 70,2 % opinó que la personalización de las notificaciones incrementaría el nivel de efectividad del sistema, mientras que un 66,7 % consideró esencial la participación activa de los usuarios para mejorar la eficacia del sistema.

Al compararlo con otros métodos para mitigar la desinformación, un 75,4 % de los participantes estimó que el método de notificaciones es más efectivo.

En resumen, los resultados indican que el prototipo posee un potencial significativo como herramienta eficaz para contrarrestar la desinformación. Este potencial reside en la capacidad del sistema para evaluar la veracidad de las noticias y enviar notificaciones, fomentando así el hábito de verificar el contenido antes de compartirlo en redes sociales. Este enfoque proactivo no solo contribuye a reducir la difusión de información engañosa, sino que también promueve la educación de los usuarios al fomentar prácticas más reflexivas y críticas al interactuar con el contenido digital.

Las evidencias estadísticas de los resultados se presentan en el apéndice C, del documento.

CAPÍTULO 6

CONCLUSIONES Y TRABAJOS FUTUROS



6.1. Introducción

En el sexto capítulo, se presentan las conclusiones y perspectivas futuras del prototipo del sistema de alerta temprana para la detección de noticias falsas en Twitter. Este representa la culminación de un extenso proceso de investigación, desarrollo y evaluación, con el objetivo de ofrecer una visión integral de los logros alcanzados y de identificar áreas de mejora y oportunidades para investigaciones futuras.

En primer lugar, se exploraron lecciones aprendidas durante el desarrollo y la evaluación del sistema. Estas incluyen la importancia de utilizar un conjunto de datos de entrenamiento diverso y representativo, la necesidad de considerar una variedad de señales para la detección de noticias falsas, y la importancia de realizar un seguimiento continuo del rendimiento del sistema para identificar áreas de mejora.

En segundo lugar, se presentan los resultados obtenidos por el prototipo del sistema. Estos demuestran que el sistema es capaz de detectar noticias falsas con un alto grado de precisión y sensibilidad.

En tercer lugar, se discute la implicación de los resultados del sistema en el contexto más amplio de la detección de desinformación en redes sociales. Los resultados sugieren que el sistema podría ser una herramienta valiosa para combatir la propagación de noticias falsas en línea.

Finalmente, se presentan las líneas de investigación que podrían potenciar y perfeccionar aún más la efectividad y aplicabilidad del sistema. Estas líneas de investigación incluyen el desarrollo de nuevos modelos de detección de noticias falsas, la integración de fuentes de información adicionales, y el desarrollo de mecanismos para la notificación efectiva de alertas a los usuarios.

De este modo, la Figura 6.1, presenta el contenido abordado en el Capítulo 6:

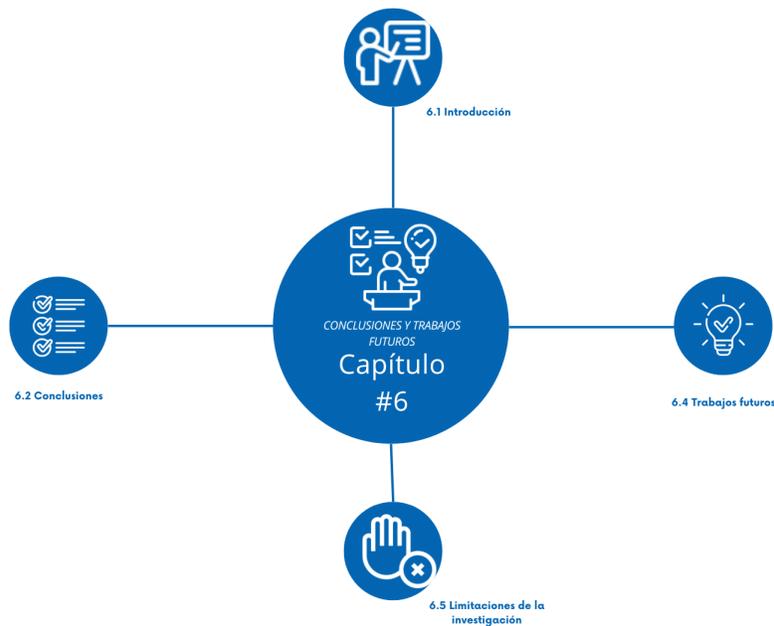


Figura 6.1: Mapa de contenido capítulo 6 (Figura propia)

6.2. Conclusiones

1. El proceso de recolección y análisis de noticias en Twitter académico permitió obtener una variedad de perspectivas sobre factores políticos, proporcionando así un panorama más completo y diverso de la información relacionada con la política en comparación con fuentes tradicionales. Sin embargo, el análisis reveló que la mayoría de las noticias relacionadas con factores políticos son falsas o engañosas, debido a que los usuarios tienden a compartir noticias sin verificarlas, facilitando la propagación de la desinformación.
2. El prototipo desarrollado fue capaz de evaluar con un alto grado de precisión la veracidad de las noticias. Esto se debe a que emplea un algoritmo que analiza las interacciones y posturas de los usuarios respecto a diferentes tweets.
3. La validación del prototipo reveló que este tiene un efecto positivo en la prevención de la desinformación y propagación de noticias falsas. Esto se debe a que ayuda a los usuarios a identificarlas y tomar decisiones informadas.
4. El diseño de un sistema de monitoreo de noticias falsas es un paso importante, pero no es suficiente para resolver el problema de la desinformación. También es necesario educar a los usuarios sobre cómo identificar noticias falsas y promover la alfabetización mediática.

5. La desinformación tiene un impacto negativo en la sociedad, porque puede conducir a la polarización política, la violencia y la toma de decisiones informadas. Es importante desarrollar soluciones para abordar la desinformación y proteger a la sociedad de sus efectos negativos.
6. Como resultado de la investigación se escribió un artículo que ofrece una visión actualizada de la problemática de las noticias falsas en el contexto político colombiano. Este, destaca las dificultades y desafíos que enfrenta el país, y ofrece recomendaciones concretas para abordarlas, en particular, detalla el proceso que se debe seguir para la elaboración de un sistema de alerta temprana, que incluye la selección adecuada de tecnologías de detección, diseño de algoritmos de clasificación eficientes y colaboración con entidades pertinentes para asegurar la efectividad del sistema. El artículo se anexa a la documentación entregada.
7. En consecuencia, el diseño de un sistema de monitoreo de noticias falsas orientado a prevenir su propagación en contextos políticos representa un paso crucial hacia la reducción de la desinformación y la polarización en la sociedad colombiana. Porque busca no solo identificar y clasificar información inexacta, sino también implementar medidas proactivas para evitar su difusión. Al abordar la desinformación en el ámbito político, se busca fortalecer la integridad del discurso público y fomentar un ambiente informado, contribuyendo así a la construcción de una sociedad más resiliente ante la manipulación de información y promoviendo la participación ciudadana basada en datos verídicos.

6.3. Limitaciones del trabajo

1. La recopilación de datos se restringió a Twitter académico, una red social dirigida a académicos y estudiantes. Esto implica que los resultados del estudio podrían carecer de generalización a otras plataformas de redes sociales, como Twitter general, Facebook, Instagram, entre otras. Las dinámicas y comportamientos específicos de la comunidad académica pueden diferir notablemente de los de usuarios en entornos más diversos, limitando la aplicabilidad de las conclusiones a un espectro más amplio de usuarios.
2. El prototipo solo se evaluó en un entorno de pruebas controlado. Esto conlleva la posibilidad de que su rendimiento no sea totalmente representativo en un entorno real, donde los usuarios pueden exhibir comportamientos más variados y espontáneos. Las condiciones controladas de la evaluación pueden no reflejar adecuadamente los desafíos y complejidades que surgen en situaciones del mundo real, lo que limita la certeza sobre su eficacia y adaptabilidad en escenarios más dinámicos.

3. La restricción en el acceso a la API de Twitter se traduce en un costo sustancial. Dado que el rendimiento óptimo del sistema depende de un extenso entrenamiento con una gran cantidad de datos, los altos costos asociados pueden representar una barrera financiera. Esta limitación no solo afecta la viabilidad económica del proyecto, sino que también restringe la capacidad de obtener una representación completa y actualizada de los datos de Twitter, impactando directamente en la calidad y exhaustividad de los resultados obtenidos.
4. La efectividad del sistema podría verse afectada por factores externos, como cambios en las políticas de privacidad de las plataformas de redes sociales o la implementación de medidas gubernamentales relacionadas con la difusión de información. Estos factores pueden influir en la accesibilidad y la disponibilidad de datos relevantes para el monitoreo de noticias falsas.

6.4. Trabajos futuros

1. Usabilidad:

- a) Desarrollar una interfaz gráfica intuitiva y fácil de usar tanto para la versión web como para la aplicación móvil.
- b) Crear una extensión o plugin que los usuarios puedan instalar en sus navegadores para recibir alertas directamente mientras navegan por Twitter.

2. Precisión:

- a) Ampliar la capacidad del sistema incorporando modelos de inteligencia artificial adicionales, como análisis de sentimientos.
- b) Aumentar la cantidad de tweets que el sistema analiza para obtener una visión más completa del panorama informativo.
- c) Refinar y mejorar los algoritmos de detección para reducir falsos positivos y negativos.

3. Cobertura:

- a) Integrar el sistema con otras plataformas de redes sociales donde también se propaga información, como Facebook, Instagram u otras redes populares.
- b) Integrar servicios de verificación de hechos e identificar sitios web conocidos por difundir información falsa.

- c) Implementar un sistema de aprendizaje continuo que se adapte y evolucione con el tiempo para mantenerse al día con las tácticas cambiantes de los propagadores de noticias falsas.
4. **Demografía:** Considerar el género de los usuarios para comprender mejor cómo las tendencias políticas se manifiestan de manera diferente en diferentes grupos demográficos

Los trabajos futuros presentados en este documento proponen mejorar el sistema de alerta temprana de noticias falsas en Twitter en tres aspectos: usabilidad, precisión y cobertura. Estas mejoras permitirían lograr una detección más efectiva de noticias falsas y, por tanto, evitar su propagación.

REFERENCIAS

- [1] J. E. F. Quinayás and J. G. M. Morcillo, “Modelo para la detección de noticias falsas en formato texto en la red social twitter, aplicado al contexto político colombiano de las elecciones presidenciales de 2022,” https://repository.icesi.edu.co/biblioteca_digital/bitstream/10906/95206/1/T02444.pdf, 2022.
- [2] M. Al-Abyadh *et al.*, “Deep sentiment analysis of twitter data using a hybrid ghost convolution neural network model,” *Computational Intelligence and Neuroscience*, vol. 2022, p. 6595799, Jul. 2022.
- [3] y. C. L. Y. Han, S. Karunasekera, “Graph neural networks with continual learning for fake news detection from social media,” *arXiv*, Aug. 2020.
- [4] H. A. y M. Gentzkow, “Social media and fake news in the 2016 election,” *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211–236, May 2017.
- [5] V. B. y A. McStay, “Fake news and the economy of emotions: Problems, causes, solutions,” *Digital Journalism*, vol. 6, pp. 1–22, Jul. 2017.
- [6] D. Zhang and V. I. Zadorozhny, “Fake news detection based on subjective opinions,” in *Advances in Databases and Information Systems*, J. Darmont, B. Novikov, and R. Wrembel, Eds., vol. 12245. Springer International Publishing, 2020, pp. 108–121.
- [7] M. Cantarella, N. Fraccaroli, and R. Volpe, “Does fake news affect voting behaviour?” *Research Policy*, vol. 52, no. 1, p. 104628, Jan. 2023.
- [8] R. M. S. Carbonell and M. del pilar Rendón Rodríguez, “El papel de las plataformas digitales y los verificadores de información en la difusión de fake news,” 2019.
- [9] C. Rodríguez-Pérez, L. Ortiz Calderón, and J. Esquivel Coronado, “Desinformación en contextos de polarización social: el paro nacional en Colombia del 21n,” *Anagramas-Rumbos y sentidos de la comunicación*, vol. 19, no. 38, pp. 129–156, Jun. 2021.

- [10] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer, “Fake news on twitter during the 2016 us presidential election,” *Science*, vol. 363, no. 6425, pp. 374–378, Jan. 2019.
- [11] “Fake news spreads faster than true news on twitter—thanks to people, not bots,” consultado el 19 de abril de 2023. [Online]. Available: <https://www.science.org/content/article/fake-news-spreads-faster-true-news-twitter-thanks-people-not-bots>
- [12] K. R. Medina, “Situación digital de colombia en 2023,” Branch Agencia, el 21 de junio de 2022, consultado el 19 de abril de 2023. [Online]. Available: <https://branch.com.co/marketing-digital/estadisticas-de-la-situacion-digital-de-colombia-en-el-2021-2022/>
- [13] R. A. E. ASALE. (2023) efectividad — diccionario de la lengua española. [Online]. Available: <https://dle.rae.es/efectividad>
- [14] N. Munier and F. Sáez, “Planificación de proyectos de i&d,” 2012, doi: 10.13140/RG.2.1.2983.0168.
- [15] S. W. J. T. y. H. L. K. Shu, A. Sliva, “Fake news detection on social media: A data mining perspective,” *arXiv*, Sep. 2017. [Online]. Available: <http://arxiv.org/abs/1708.01967>
- [16] D. M. J. L. et al., “The science of fake news,” *Science*, vol. 359, no. 6380, pp. 1094–1096, Mar. 2018.
- [17] S. B. y. T. S.-E. T. Quandt, L. Frischlich, “Fake news,” pp. 1–6, Apr. 2019.
- [18] D. E. D. M. y. M. M. B. F. Monti, F. Frasca, “Fake news detection on social media using geometric deep learning,” *arXiv*, Feb. 2019.
- [19] y. Y. L. N. Ruchansky, S. Seo, “Csi: A hybrid deep model for fake news detection,” in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, Nov. 2017, pp. 797–806.
- [20] S. W. D. L. y. H. L. K. Shu, L. Cui, “defend: Explainable fake news detection,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining, KDD '19*. New York, NY, USA: Association for Computing Machinery, Jul. 2019, pp. 395–405.
- [21] A. G. W. Q. y. M. S. M. Cinelli, G. De Francisci Morales, “The echo chamber effect on social media,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 9, p. e2023301118, Mar. 2021.

- [22] y. F. B. S. I. O. Furman, K. B. Gürel, “‘as reliable as a kalashnikov rifle’: How sputnik news promotes russian vaccine technologies in the turkish twittersphere,” *Social Media + Society*, vol. 9, no. 1, p. 20563051221150418, Jan. 2023.
- [23] M. D. V. et al., “Echo chambers in the age of misinformation,” *arXiv*, Dec. 2015.
- [24] —, “The spreading of misinformation online,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 3, pp. 554–559, Jan. 2016.
- [25] J. L. E. y S. Lecheler, “Fake news as a two-dimensional phenomenon: a framework and research agenda,” *Annals of the International Communication Association*, vol. 43, no. 2, pp. 97–116, Apr. 2019.
- [26] B. McNair, *Fake News: Falsehood, Fabrication and Fantasy in Journalism*. Routledge, 2017.
- [27] “Defining ‘fake news’: A typology of scholarly definitions: Digital journalism: Vol 6, no 2,” Jan. 2024. [Online]. Available: <https://www.tandfonline.com/doi/abs/10.1080/21670811.2017.1360143>
- [28] “Fake news detection in social media — semantic scholar,” Jan. 2024. [Online]. Available: <https://www.semanticscholar.org/paper/Fake-news-detection-in-social-media-Stahl/b202b4b7124b774391109dc47a33e17224b12295>
- [29] N. R. Hanson, “A note on statements of fact,” *Analysis*, vol. 13, no. 1, p. 24, Oct. 1952.
- [30] F. Pierri and S. Ceri, “False news on social media: A data-driven survey,” *arXiv*, Jan 2020.
- [31] D. R. S. Vosoughi and S. Aral, “The spread of true and false news online,” *Science*, vol. 359, p. 1146–1151, Mar 2018.
- [32] N. Kshetri and J. Voas, “The economics of ‘fake news’,” *IT Professional*, vol. 19, no. 6, p. 8–12, Nov 2017.
- [33] E. J. Fox and S. J. Hoch, “Cherry-picking,” *Journal of Marketing*, vol. 69, no. 1, p. 46–62, Jan 2005.
- [34] “Detection and resolution of rumours in social media: A survey: Acm computing surveys: Vol 51, no 2,” 2018. [Online]. Available: <https://dl.acm.org/doi/10.1145/3161603>

- [35] M. D. V. et al., “The spreading of misinformation online,” *Proc Natl Acad Sci U S A*, vol. 113, no. 3, p. 554–559, Jan 2016. [Online]. Available: <https://doi.org/10.1073/pnas.1517441113>
- [36] F. Allen and D. Gale, “Stock-price manipulation,” *The Review of Financial Studies*, vol. 5, no. 3, p. 503–529, Jul 1992. [Online]. Available: <https://doi.org/10.1093/rfs/5.3.503>
- [37] B. Hofmann, “Fake facts and alternative truths in medical research,” *BMC Med Ethics*, vol. 19, no. 1, p. 4, Jan 2018. [Online]. Available: <https://doi.org/10.1186/s12910-018-0243-z>
- [38] D. H. H. T. Phan, N. T. Nguyen, “Fake news detection: A survey of graph neural network methods,” *Applied Soft Computing*, vol. 139, p. 110235, May 2023.
- [39] “Utility - meaning, types measures,” *VEDANTU*, 2024. [Online]. Available: <https://www.vedantu.com/commerce/utility>
- [40] “(pdf) what is a prototype? what are the roles of prototypes in companies?” 2024. [Online]. Available: https://www.researchgate.net/publication/323198004_What_is_a_Prototype_What_are_the_Roles_of_Prototypes_in_Companies
- [41] “Ux daily: The world’s largest open-source ux design library,” *The Interaction Design Foundation*, 2024. [Online]. Available: <https://www.interaction-design.org/literature/article/overview>
- [42] “caracterizar — diccionario de la lengua española,” *Diccionario de la lengua española - Edición del Tricentenario*, 2024. [Online]. Available: <https://dle.rae.es/caracterizar>
- [43] C. M. V. N. E. Gallo Restrepo, Y. A. Meneses Copete, “Caracterización poblacional vista desde la perspectiva del desarrollo humano y el enfoque diferencial,” *Investigación y Desarrollo*, vol. 22, no. 2, p. 360–401, Jul 2014.
- [44] A. Komninos, “An introduction to usability,” *The Interaction Design Foundation*, 2024. [Online]. Available: <https://www.interaction-design.org/literature/article/an-introduction-to-usability>
- [45] A. M. Lund and D. Wixon, “Usability testing of computer interfaces,” in *Human-computer interaction: Software and hardware interfaces*, G. Salvendy, Ed. Elsevier, 1985, pp. 315–327, consultado el 19 de abril de 2023.
- [46] B. A. Kitchenham, D. Budgen, and P. O. Brereton, “The value of mapping studies: a participantobserver case study,” in *Proceedings of the 14th international conference on Evaluation and Assessment in Software Engineering*. BCS Learning Development Ltd., 2010, pp. 25–33.

- [47] I. Nadeem, D. Li, and Kanwa, "EfnD: A semantic, visual, and socially augmented deep framework for extreme fake news detection," *Sustainability*, vol. 15, p. 1133, Dec. 2022.
- [48] K. L. Gwebu, J. Wang, and E. Zifla, "Can warnings curb the spread of fake news? the interplay between warning, trust and confirmation bias," *Behaviour & Information Technology*, vol. 41, no. 16, pp. 3552–3573, Dec. 2022.
- [49] N. C. Wickramarathna, T. D. Jayasiriwardena, M. Wijesekara, P. B. Munasinghe, and G. U. Ganegoda, "A framework to detect twitter platform manipulation and computational propaganda," in *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*, Nov. 2020, pp. 214–219.
- [50] J. Jussila, A. Suominen, A. Partanen, and T. Honkanen, "Text analysis methods for misinformation-related research on Finnish language twitter," *Future Internet*, vol. 13, p. 157, Jun. 2021.
- [51] S. Sharma, M. Saraswat, and A. K. Dubey, "Fake news detection on twitter," *International Journal of Web Information Systems*, vol. 18, no. 5–6, pp. 388–412, Dec. 2022.
- [52] M. P. Thilakarathna *et al.*, "Hybrid approach and architecture to detect fake news on twitter in real-time using neural networks," in *2020 5th International Conference on Information Technology Research (ICITR)*, Dec. 2020, pp. 1–6.
- [53] S. Patel and I. U. Hassan, "Identifying fake news with various machine learning model," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Sep. 2021, pp. 1–5.
- [54] (2014) Facilitating twitter data analytics: Platform, tools, and infrastructure. Consultado: el 16 de enero de 2024. [Online]. Available: https://ceur-ws.org/Vol-838/paper_08.pdf
- [55] K. Tao and C. Hauff. (2014) Facilitating twitter data analytics: Platform, tools, and infrastructure. Consultado: el 16 de enero de 2024. [Online]. Available: <https://www.semanticscholar.org/paper/Facilitating-Twitter-data-analytics%3A-Platform%2C-and-Tao-Hauff/74403e3971c70286444397e50d567fd6941a1b4d>
- [56] Y. Takemura and A. Tajima. (2017) Tweet classification based on their lifetime. Consultado: el 16 de enero de 2024. [Online]. Available: <https://www.semanticscholar.org/paper/Tweet-classification-based-on-their-lifetime-Takemura-Tajima/7e94184ea8943d3be7ae1143d2d62784913df5a2>

- [57] Lesjak. (2024) Usuarios de twitter por edad. Consultado: el 16 de enero de 2024. [Online]. Available: <https://tridenstechnology.com/es/twitter-usuarios-estadisticas/>
- [58] (2024) Redes sociales y meta: porcentaje de usuarios por edad en colombia en 2023. Consultado: el 16 de enero de 2024. [Online]. Available: <https://es.statista.com/estadisticas/1367608/redes-sociales-porcentaje-de-usuariose-de-meta-por-edad-en-colombia/>
- [59] S. X. G. Galvis, D. R. G. Tavera, and J. F. A. Pongutá, "El lenguaje político en twitter durante la segunda vuelta presidencial colombia 2018," *Anagramas -Rumbos y sentidos de la comunicación-*, vol. 20, no. 39, pp. 107–127, Dec. 2021.
- [60] (2017) Adaptación al español del cuestionario de usabilidad de sistemas informáticos (csuq): Spanish language adaptation of the computer systems usability questionnaire (csuq). Consultado: el 16 de enero de 2024. [Online]. Available: https://www.researchgate.net/publication/312962078_Adaptacion_al_espanol_del_Cuestionario_de_Usabilidad_de_Sistemas_Informaticos_CSUQ_Spanish_language_adaptation_of_the_Computer_Systems_Usability_Questionnaire_CSUQ



Apéndice A

CONSENTIMIENTO INFORMADO PARA LA REALIZACIÓN DE PRUEBAS



En este apéndice, se proporciona una descripción detallada del proceso de obtención del consentimiento informado necesario para llevar a cabo las pruebas:

FORMATO DE CONSENTIMIENTO INFORMADO PARA UN PROYECTO DE INVESTIGACIÓN

El Trabajo de Grado titulado **Sistema de alerta temprana para monitorear la propagación de noticias falsas: Caso de estudio contexto político colombiano**, cuyo objetivo es **Diseñar un sistema de monitoreo de noticias falsas (fake news) que prevenga su propagación en contextos políticos, con el fin de disminuir la desinformación y polarización de la sociedad en Colombia.**, el cual se desarrollará por los estudiantes Laura Sofía Daza Rosero y César Enrique Manzano Velasco, pertenecientes al programa de Ingeniería Electrónica y Telecomunicaciones de la Facultad de Ingeniería Electrónica y Telecomunicaciones de la Universidad del Cauca, bajo la dirección de César Alberto Collazos Ordóñez adscrito al Departamento de Sistemas, se realizará como requisito para optar al título de Ingeniera en Electrónica y Telecomunicaciones.

1. **Participantes:** Personas en el rango de edad de 18 a 35 años.
2. **Propósito de la investigación:** El propósito central de esta investigación se enmarca en la validación del prototipo desarrollado, orientada a analizar de manera integral sus repercusiones en la prevención activa de la desinformación y la diseminación de noticias falsas. En este contexto, la efectividad del prototipo se evalúa minuciosamente, no solo desde una perspectiva técnica, sino principalmente en términos del comportamiento observado en los usuarios al enfrentarse a información potencialmente engañosa. La evaluación se extiende a la capacidad del sistema para lograr el efecto deseado en la contención y contrarresto de la desinformación, considerando las reacciones y decisiones de los usuarios ante la exposición a noticias, destacando su papel clave en la prevención activa y en la construcción de una percepción más precisa y fiable del entorno informativo en línea. De este modo, la investigación no solo busca validar la funcionalidad técnica del prototipo, sino también entender su impacto real en la conducta de los usuarios y, por ende, en la salvaguarda de la integridad informativa. Tipo de intervención de la investigación: Participará activamente en la interacción con un conjunto de noticias (falsas o verdaderas) y, al finalizar, expresará su opinión y evaluación sobre el sistema de alerta al que se haya enfrentado..
3. **Selección de las participantes:** Usted está cordialmente invitado a formar parte de este proyecto de investigación por estar en un rango de personas entre 15 y 35 años.
4. **Participación voluntaria:** Usted puede elegir si quiere ser parte del proyecto de investigación. Si no quiere ser parte del estudio su trabajo continuará

y nada cambiará. Aún si usted acepta ser parte del estudio ahora, usted se puede arrepentir luego y dejar de participar.

5. **Procedimientos:** Si Usted acepta participar en el estudio: se le presentarán algunas noticias con el fin de evaluar su reacción ante un conjunto de información. Posteriormente, se llevará a cabo una encuesta para validar su comportamiento en relación con el sistema.
6. **Riesgos y molestias:** Usted debe colaborar con su asistencia en la entrevista en profundidad. Esta será realizada por un estudiante del programa de Ingeniería Electrónica y Telecomunicaciones de la Universidad del Cauca.
7. **Beneficios:** Al concluir el proyecto de investigación, los descubrimientos obtenidos se presentan como valiosos aportes que pueden ser instrumentalizados para el avance en el desarrollo de sistemas más efectivos. Estos sistemas, al integrar los conocimientos adquiridos, estarán mejor equipados para contribuir de manera significativa a la mitigación y control de la difusión de noticias falsas. La aplicación práctica de estos hallazgos se vislumbra como una medida proactiva en la lucha contra la desinformación, fortaleciendo las bases para la construcción de plataformas más robustas y eficientes en la detección y gestión de contenidos engañosos. Este enfoque no solo respalda la mejora continua de la tecnología en esta área crítica, sino que también promueve la capacidad de adaptación de los sistemas ante las cambiantes tácticas de propagación de noticias falsas, consolidando así un entorno más resistente y confiable en la esfera informativa.
8. **Incentivos:** Usted no recibirá ningún incentivo de índole económico (no se le dará dinero) por participar en el estudio.
9. **Confidencialidad:** La información derivada de la investigación será manejada de manera exclusiva por nuestro grupo de investigación. La información será guardada de manera segura por parte del equipo. Nadie fuera de nuestro equipo de investigación verá la información sobre usted.

Cuando expliquemos la investigación a otras personas, no usaremos su nombre o nada que permita que otras personas conozcan su identidad. La información será guardada, cumpliendo los criterios de confidencialidad y respeto. Cabe resaltar que todos los datos personales utilizados en este proyecto no serán utilizados en otras investigaciones. Se realizará un análisis estadístico, al cual solo tendrá acceso el grupo de investigación supervisado por el Ing.César Alberto Collazos Ordóñez, de la Facultad de Ingeniería Electrónica de la Universidad del Cauca, Campus de Tulcán. Celular: +57 311 308 1728.
10. **Divulgación de resultados:** Una vez toda la información haya sido analizada, escribiremos sobre nuestros resultados sin mencionar los datos de los

participantes.

11. **Derecho a rehusar o a retirarse:** Como se dijo anteriormente, su participación es voluntaria. Usted puede retirarse del estudio en cualquier momento si así lo desea. En ese caso, su información será eliminada. Esta investigación contiene los elementos éticos que la ley y la doctrina exigen (código *Helsinki* - código *Nuremberg* resolución 008430 de 1993) que rigen la ética en la investigación científica en Colombia. Se garantiza total confidencialidad con los datos recolectados.
12. **Información de contacto:** Si tiene preguntas, las puede hacer ahora o posteriormente. Se le dará una copia escrita de este consentimiento. Si tiene preguntas adicionales, por favor contáctenos a través de el Dr. César Alberto Collazos Ordóñez. Celular: +57 311 308 1728.
13. **Certificado de consentimiento:** Entiendo que se me va a realizar una encuesta para evaluar la efectividad de un sistema de alerta temprana. Entiendo que no existe ningún riesgo. Sé que no recibiré dinero, sino el beneficio que los resultados de investigación que ayuden el desarrollo y mejora de nuevos sistemas que permitan mitigar la propagación de noticias falsas. Se me ha dado el nombre y dirección de un investigador que puede ser contactado fácilmente.

He leído o me ha sido leída la información precedente. He tenido la oportunidad de hacer preguntas. Estoy satisfecho/a con las respuestas a todas mis preguntas. Doy consentimiento voluntario para hacer parte en este estudio. También puedo retirarme en cualquier momento.

Nombre legible del participante
Cédula.

He leído exactamente o he sido testigo de la lectura correcta del consentimiento al participante potencial, y éste ha tenido la posibilidad de hacer preguntas. Confirmando que el/la participante ha dado consentimiento libremente.

Laura Sofía Daza Rosero
Estudiante. CC:1.002'972.786
Código: 100617020743

César Enrique Manzano Velasco
Estudiante. CC:1.002'971.309
Código:100617020756

César Alberto Collazos
Director.

En constancia de aceptación, se firma el Acta por los que en ella intervienen, a los ___ días del mes de _____ del 20__ y se da una copia de este consentimiento informado cada participante.

Apéndice B

ENCUESTA CSUQ



En este apéndice se proporciona la encuesta CSUQ, herramienta diseñada para evaluar la calidad de uso de un producto de software.

CUESTIONARIO DE USABILIDAD EN SISTEMAS INFORMÁTICOS (CSUQ)

	Totalmente en				Totalmente de		
	desacuerdo				acuerdo		
	1	2	3	4	5	6	7
1 En general, estoy satisfecho con lo fácil que es utilizar este sitio web.	<input type="radio"/>						
2 Fue simple usar este sitio web.	<input type="radio"/>						
3 Soy capaz de completar mi trabajo rápidamente utilizando este sitio web.	<input type="radio"/>						
4 Me siento cómodo utilizando este sitio web.	<input type="radio"/>						
5 Fue fácil aprender a utilizar este sitio web.	<input type="radio"/>						
6 Creo que me volví experto rápidamente utilizando este sitio web.	<input type="radio"/>						
7 El sitio web muestra mensajes de error que me dicen claramente cómo resolver los problemas.	<input type="radio"/>						
8 Cada vez que cometo un error utilizando el sitio web, lo resuelvo fácil y rápidamente	<input type="radio"/>						
9 La información (como ayuda en línea, mensajes en pantalla y otra documentación) que provee este sitio web es clara.	<input type="radio"/>						
10 Es fácil encontrar en el sitio web la información que necesito.	<input type="radio"/>						
11 La información que proporciona el sitio web fue efectiva ayudándome a completar las tareas.	<input type="radio"/>						
12 La organización de la información del sitio web en la pantalla fue clara.	<input type="radio"/>						
13 La interfaz del sitio web fue placentera.	<input type="radio"/>						
14 Me gustó utilizar el sitio web.	<input type="radio"/>						
15 El sitio web tuvo todas las herramientas que esperaba que tuviera.	<input type="radio"/>						
16 En general, estuve satisfecho con el sitio web.	<input type="radio"/>						

Figura B.1: Cuestionario CSUQ

Apéndice C

RESULTADOS GRÁFICOS DE LAS ENCUESTAS



En este apéndice, se exhiben los resultados gráficos de la encuesta:

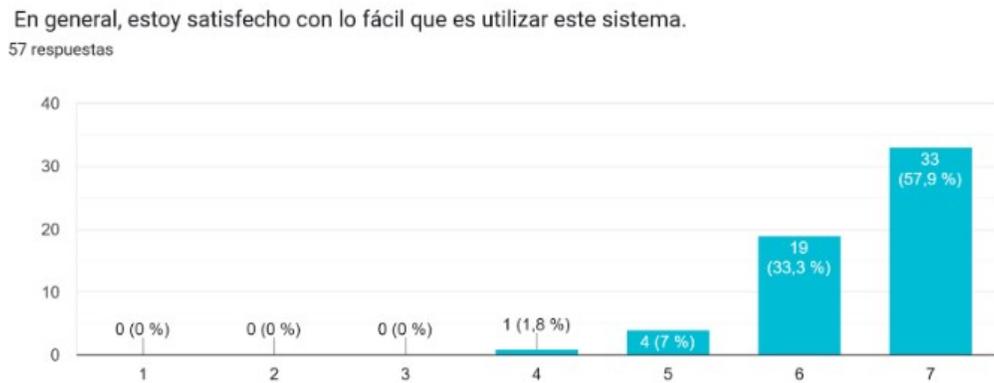


Figura C.1: Gráfica de resultados para la afirmación "En general, estoy satisfecho con lo fácil que es utilizar este sistema"(Figura propia)

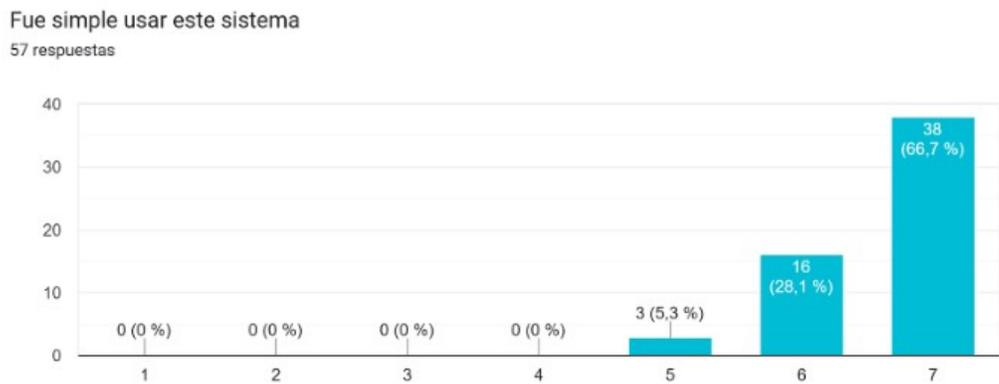


Figura C.2: Gráfica de resultados para la afirmación "Fue simple usar este sistema"(Figura propia)

Me siento cómodo utilizando este sistema.

57 respuestas

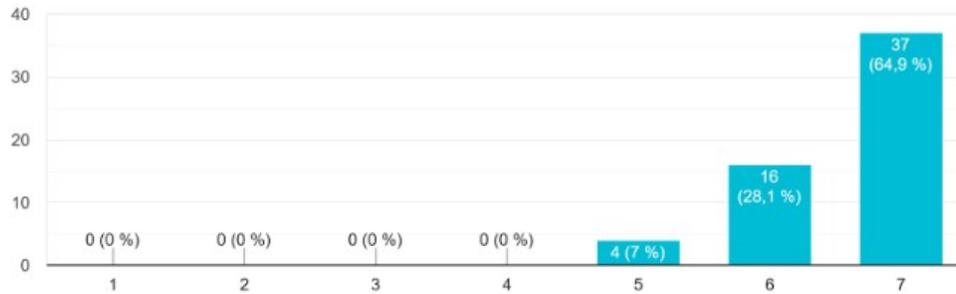


Figura C.3: Gráfica de resultados para la afirmación "Me siento cómodo utilizando este sistema"(Figura propia)

Fue fácil aprender a utilizar este sistema.

57 respuestas

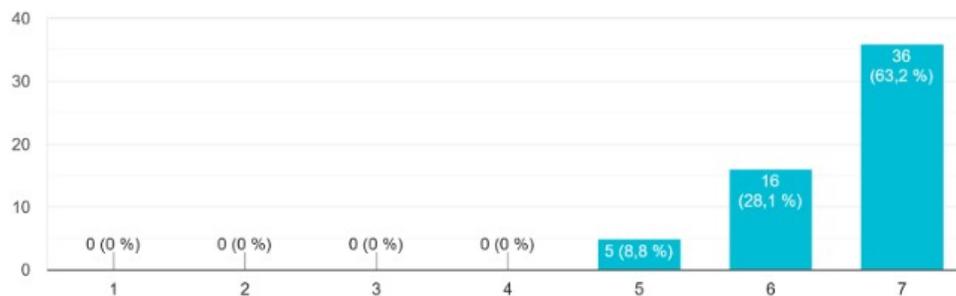


Figura C.4: Gráfica de resultados para la afirmación "Fue fácil aprender a utilizar este sistema"(Figura propia)

Creo que entendí rápidamente la finalidad de este sistema.

57 respuestas

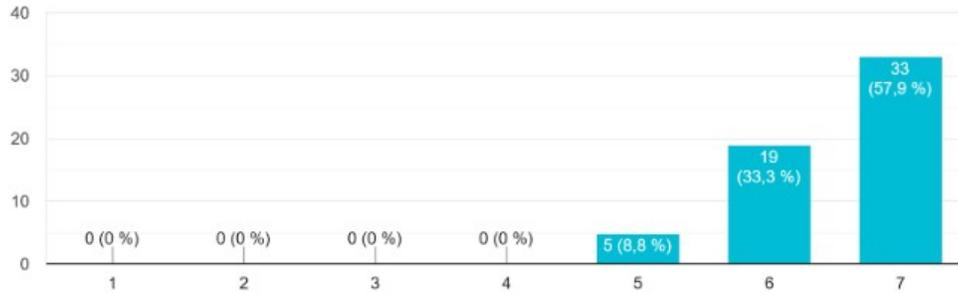


Figura C.5: Gráfica de resultados para la afirmación "Creo que entendí rápidamente la finalidad de este sistema"(Figura propia)

La información que proporciona el sistema fue efectiva ayudándome a completar las tareas.

57 respuestas

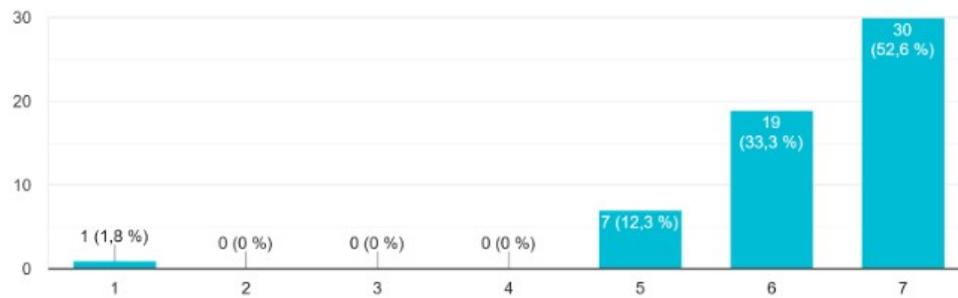


Figura C.6: Gráfica de resultados para la afirmación "La información que proporciona el sistema fue efectiva ayudándome a completar las tareas"(Figura propia)

La interfaz del sistema fue placentera.

57 respuestas

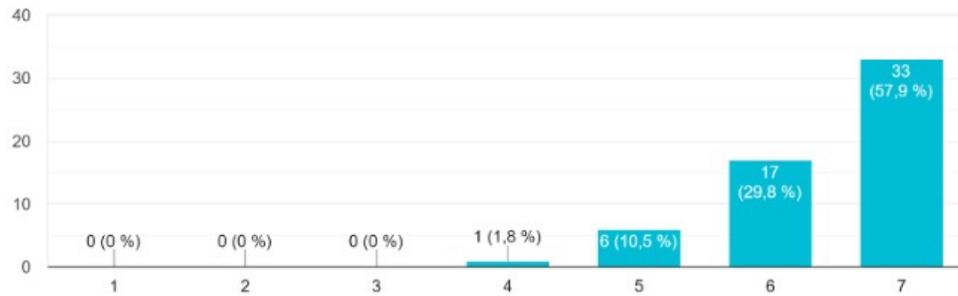


Figura C.7: Gráfica de resultados para la afirmación "La interfaz del sistema fue placentera"(Figura propia)

Me gustó utilizar el sistema.

57 respuestas

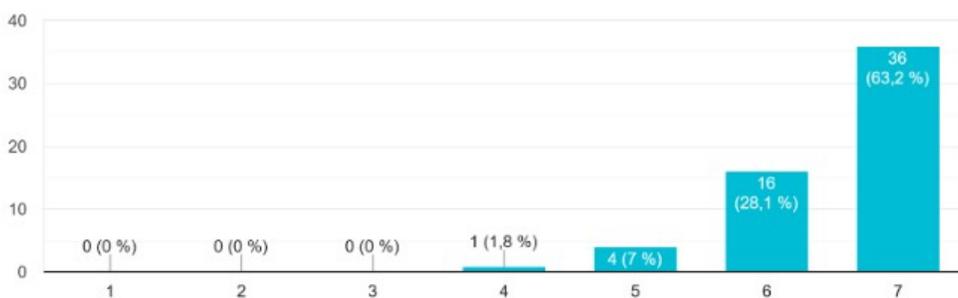


Figura C.8: Gráfica de resultados para la afirmación "Me gustó utilizar el sistema"(Figura propia)

El sistema tuvo todas las herramientas que esperaba que tuviera.

57 respuestas

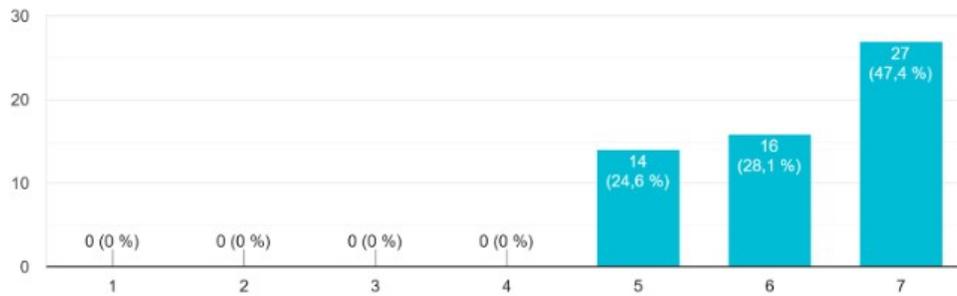


Figura C.9: Gráfica de resultados para la afirmación " El sistema tuvo todas las herramientas que esperaba que tuviera"(Figura propia)

En general, estuve satisfecho con el sistema.

57 respuestas

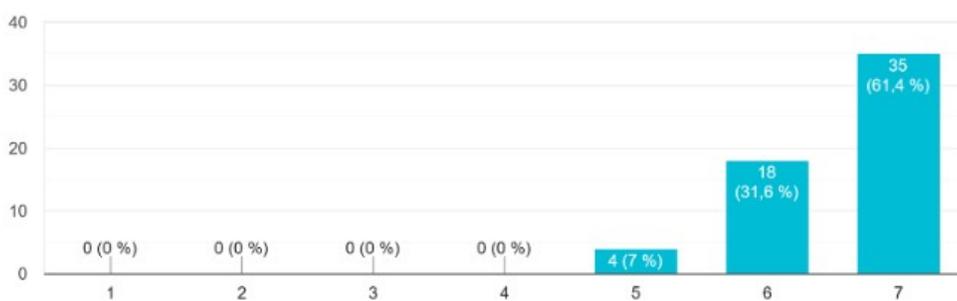


Figura C.10: Gráfica de resultados para la afirmación " En general, estuve satisfecho con el sistema"(Figura propia)

¿Cómo calificaría su capacidad para identificar noticias falsas antes de usar el sistema?

57 respuestas

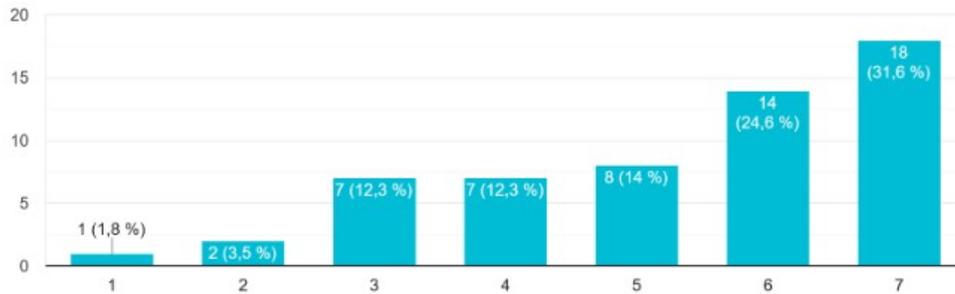


Figura C.11: Gráfica de resultados para la pregunta "¿Cómo calificaría su capacidad para identificar noticias falsas antes de usar el sistema?" (Figura propia)

¿Notó alguna mejora en su capacidad para identificar noticias falsas después de usar el sistema de alerta temprana?

57 respuestas

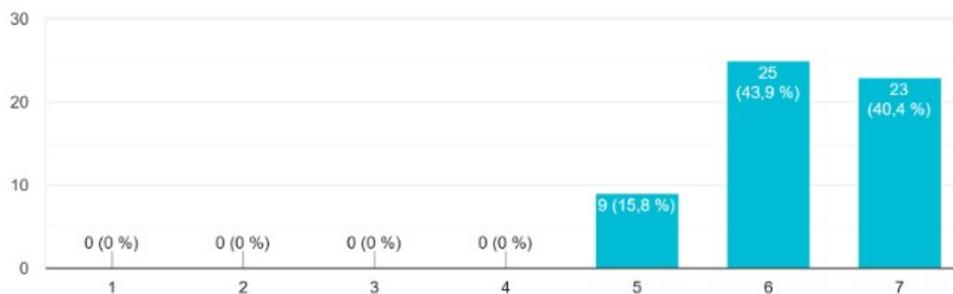


Figura C.12: Gráfica de resultados para la pregunta "¿Notó alguna mejora en su capacidad para identificar noticias falsas después de usar el sistema de alerta temprana?" (Figura propia)

¿Influyó el sistema en su motivación para buscar más información sobre las noticias después de recibir las notificaciones, especialmente en términos de verificar la veracidad de las noticias?

57 respuestas

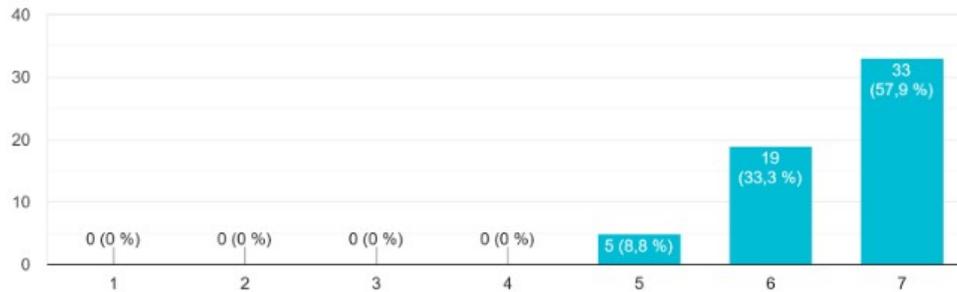


Figura C.13: Gráfica de resultados para la pregunta "¿Influyó el sistema en su motivación para buscar más información sobre las noticias después de recibir las notificaciones, especialmente en términos de verificar la veracidad de las noticias?"(Figura propia)

¿Hubo alguna noticia que inicialmente creyó que era falsa y luego el sistema identificó como verdadera (Visceversa)?

57 respuestas

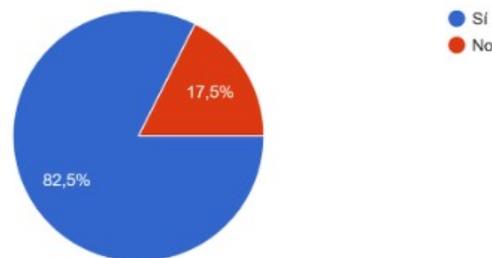


Figura C.14: Gráfica de resultados para la pregunta "¿Hubo alguna noticia que inicialmente creyó que era falsa y luego el sistema identificó como verdadera (Visceversa)?"(Figura propia)

¿Cómo percibe la rapidez de respuesta de un sistema de alerta temprana en la detección de desinformación?

57 respuestas

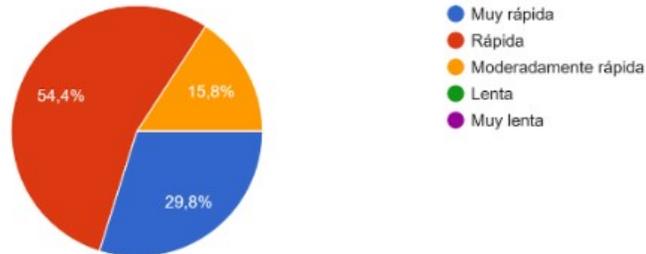


Figura C.15: Gráfica de resultados para la pregunta "¿Cómo percibe la rapidez de respuesta de un sistema de alerta temprana en la detección de desinformación?"(Figura propia)

¿Cree que la personalización de las notificaciones, adaptadas a sus intereses, aumentaría la efectividad del sistema de alerta temprana?

57 respuestas

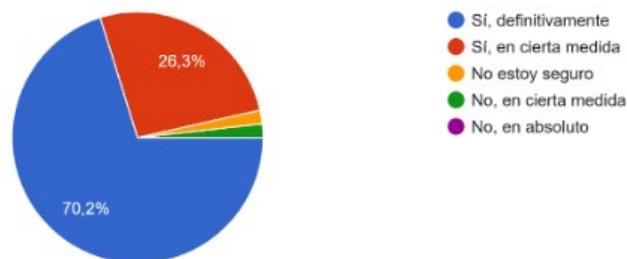


Figura C.16: Gráfica de resultados para la pregunta "¿Cree que la personalización de las notificaciones, adaptadas a sus intereses, aumentaría la efectividad del sistema de alerta temprana?"(Figura propia)

¿Considera que la participación activa de la comunidad de usuarios es esencial para mejorar la eficacia del sistema de alerta temprana?

57 respuestas

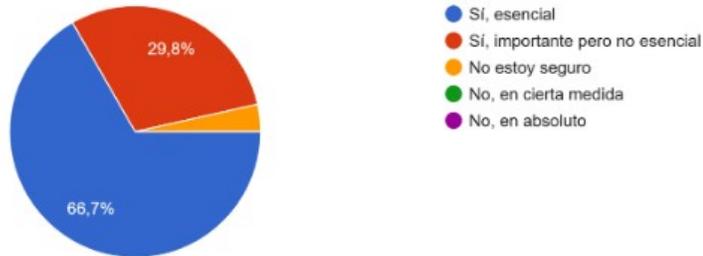


Figura C.17: Gráfica de resultados para la pregunta "¿Considera que la participación activa de la comunidad de usuarios es esencial para mejorar la eficacia del sistema de alerta temprana?"(Figura propia)

¿Opina que la transparencia en el funcionamiento del sistema de alerta temprana aumentaría su confianza en sus resultados?

57 respuestas

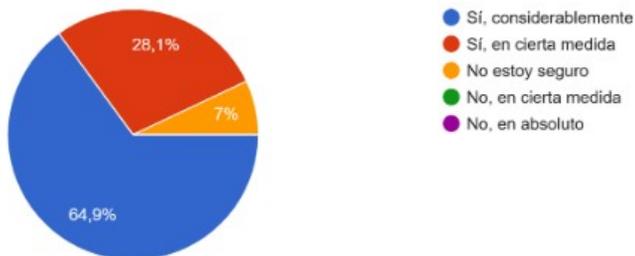


Figura C.18: Gráfica de resultados para la pregunta "¿Opina que la transparencia en el funcionamiento del sistema de alerta temprana aumentaría su confianza en sus resultados?"(Figura propia)

¿Cómo evalúa la efectividad del sistema de alerta temprana en comparación con otros métodos de combatir la desinformación?

57 respuestas

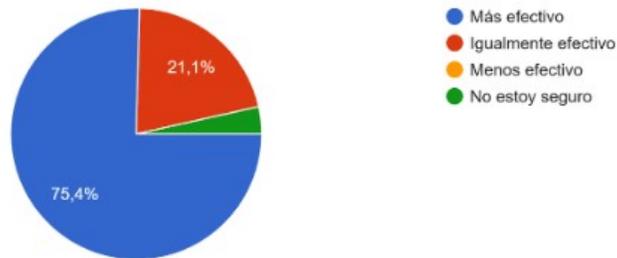


Figura C.19: Gráfica de resultados para la pregunta "¿Cómo evalúa la efectividad del sistema de alerta temprana en comparación con otros métodos de combatir la desinformación?"(Figura propia)

¿Hasta qué punto le resulta efectivo un sistema de alerta temprana en la lucha contra la desinformación mediante la implementación de notificaciones?

57 respuestas

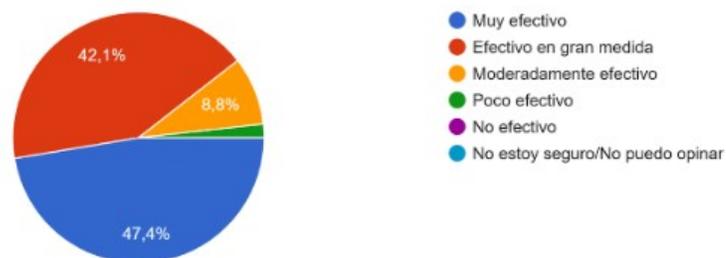


Figura C.20: Gráfica de resultados para la pregunta "¿Hasta qué punto le resulta efectivo un sistema de alerta temprana en la lucha contra la desinformación mediante la implementación de notificaciones?"(Figura propia)

¿En qué medida confía en la capacidad de un sistema de alerta temprana para detectar y combatir la desinformación?

57 respuestas

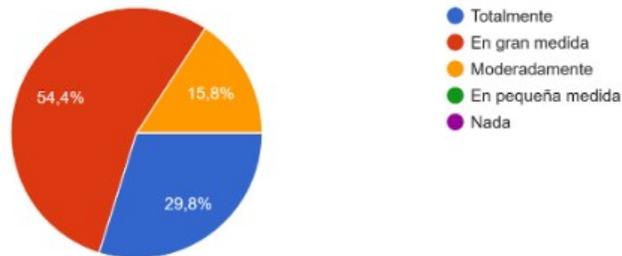


Figura C.21: Gráfica de resultados para la pregunta "¿En qué medida confía en la capacidad de un sistema de alerta temprana para detectar y combatir la desinformación?"(Figura propia)

¿Considera que la rapidez de las notificaciones afecta la eficacia del sistema de alerta temprana en combatir la desinformación?

57 respuestas

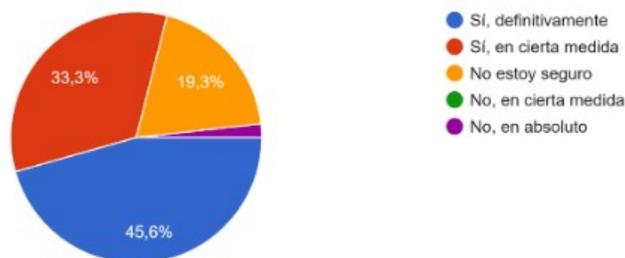


Figura C.22: Gráfica de resultados para la pregunta "¿Considera que la rapidez de las notificaciones afecta la eficacia del sistema de alerta temprana en combatir la desinformación?"(Figura propia)

¿Cree que la implementación de notificaciones puede influir en su percepción y comprensión de la información?

57 respuestas

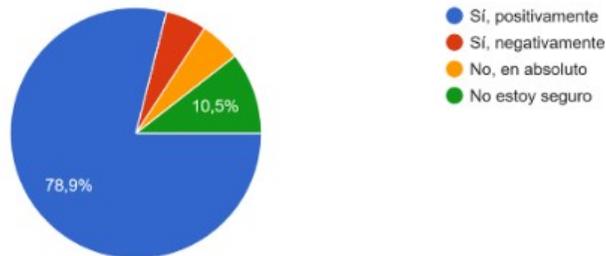


Figura C.23: Gráfica de resultados para la pregunta "¿Cree que la implementación de notificaciones puede influir en su percepción y comprensión de la información?"(Figura propia)

¿Cuál sería la frecuencia ideal de notificaciones para mantener informado al usuario y combatir la desinformación de manera efectiva?

57 respuestas

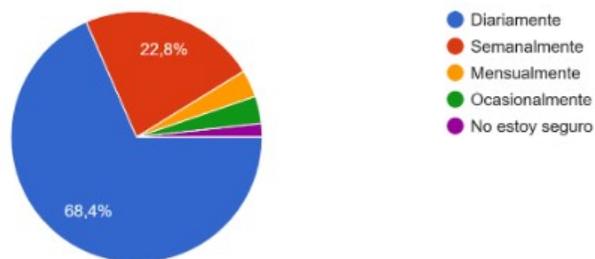


Figura C.24: Gráfica de resultados para la pregunta "¿Cuál sería la frecuencia ideal de notificaciones para mantener informado al usuario y combatir la desinformación de manera efectiva?"(Figura propia)

¿Piensa que la efectividad del sistema de alerta temprana dependería de la colaboración entre plataformas, usuarios y autoridades?

57 respuestas

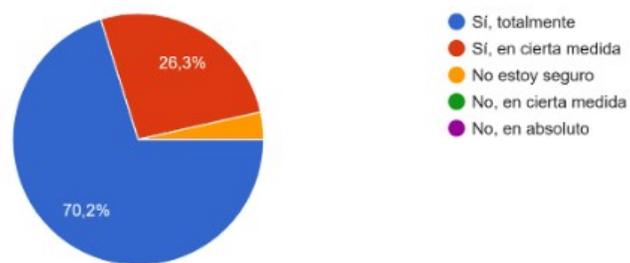


Figura C.25: Gráfica de resultados para la pregunta "¿Piensa que la efectividad del sistema de alerta temprana dependería de la colaboración entre plataformas, usuarios y autoridades?"(Figura propia)