



Formando líderes para la construcción
de un nuevo país en paz

Aplicación web con IA basada en NLP para detectar noticias falsas, utilizando Scrum y principios de arquitectura por capas para una implementación ágil y efectiva.

Samuel Andres Celis Lizcano

1091964042

Informe final del curso Ingeniería de Software III

Universidad de Pamplona

Facultad de Ingenierías y Arquitectura

Ingeniería de Sistemas

Sede Villa del Rosario

2025



Tabla de contenido

1.	Resumen.....	5
2.	Abstract	6
3.	Introducción	7
4.	Pregunta de Investigación	9
5.	Justificación	12
6.	Objetivos	14
7.	Estado del arte.....	15
7.1	Marco Referencial.....	15
7.1.1	Antecedentes Internacionales.....	15
7.1.2	Antecedentes Nacionales	15
7.1.3	Antecedentes Regionales	16
7.2	Marco Conceptual.....	17
7.3	Marco Contextual.....	18
7.3.1	Contexto social y comunicacional	18
7.3.2	Contexto tecnológico y de desarrollo.....	19
7.3.3	Contexto académico y metodológico.....	19
7.4	Marco Legal	20
8.	Metodología	20
8.1	Desde un punto de vista de investigación teórico y filosófico	20
8.2	Desde el punto de vista de desarrollo.....	21
9.	Selección del patrón de diseño MVC.....	22
9.1	Separación de responsabilidades.....	23
9.2	Escalabilidad y mantenimiento	24



9.3	Fácil integración con frameworks y tecnologías modernas	24
9.4	Experiencia de usuario mejorada	25
9.5	Conclusión	25
10.	Desarrollo.....	25
10.1	Objetivo específico I	25
10.1.1	Herramientas	26
10.1.2	Tecnologías	28
10.2	Objetivo Especifico II	35
10.2.1	Selección de las Herramientas	35
10.2.2	Selección de las Tecnologías	38
10.3	Objetivo Especifico III.....	41
10.3.1	Front-End	42
10.3.2	Back-End.....	45
10.3.3	Modelo de Clasificación	46
10.4	Objetivo Especifico IV.....	50
11.	Conclusiones y recomendaciones	53
12.	Referencias.....	54

Tabla de Ilustraciones

Ilustración 1	Fases de la Metodología SCRUM.....	8
Ilustración 2	Fases de la metodología SCRUM	21
Ilustración 3	Relación objetivos con las fases de la metodología	22

"Formando líderes para la construcción de un nuevo país en paz"



Ilustración 4 Modelo Vista Controlador (MCV)	23
Ilustración 5 Interfaz Gráfica de Detector Fake News	42
Ilustración 6 Pagina inicial	43
Ilustración 7 Estructura de carpetas Front-End	43
Ilustración 8 Archivo app.tsx	44
Ilustración 9 Componente Enlaces	45
Ilustración 10 Código Back-End	46
Ilustración 11 Composición principal del Modelo de Clasificación	47
Ilustración 12 Composición secundaria del Modelo de Clasificación	48
Ilustración 13 Composición final del Modelo de Clasificación	49
Ilustración 14 Prueba 1	50
Ilustración 15 Prueba 2	51
Ilustración 16 Prueba 3	51
Ilustración 17 Prueba 4	52

Tabla de Tablas

Tabla 1 Comparación de Herramientas de versiones	35
Tabla 2 Comparación de IDEs con tecnologías	36
Tabla 3 Comparación este los Software de manejo de SCRUM	37
Tabla 4 Fortalezas de cada tecnología de desarrollo	38
Tabla 5 Comparación de las fortalezas de cada tecnología para el Back-End	39
Tabla 6 Comparación de Modelos y librerías de Clasificación	40



1. Resumen

El presente proyecto tiene como objetivo desarrollar una aplicación web con inteligencia artificial (IA) basada en procesamiento de lenguaje natural (NLP) para detectar noticias falsas. Su importancia radica en combatir la desinformación que circula en redes sociales, promoviendo el acceso a información veraz y fortaleciendo la confianza en los medios digitales. Las noticias serán obtenidas de fuentes públicas en línea, como portales informativos, redes sociales y bases de datos de noticias verificadas, para alimentar y entrenar los modelos de detección. Para su desarrollo, se empleará la metodología ágil **Scrum**, que permite una gestión flexible, colaborativa y eficiente. En la **Planificación del Sprint** se definirán funcionalidades clave como el análisis de textos y la clasificación automática. Durante la fase de **Desarrollo**, estas tareas se implementarán en ciclos cortos. En el **Daily Scrum**, el equipo verificará avances y resolverá bloqueos. Al finalizar cada ciclo, se realizará una **Revisión del Sprint** para evaluar los resultados y una **Retrospectiva** para proponer mejoras continuas.

El sistema se diseñará bajo una **arquitectura por capas**, organizando la solución en capas de presentación, lógica de negocio y acceso a datos, lo que favorece la escalabilidad y el mantenimiento. Asimismo, se aplicará el patrón de diseño **Modelo-Vista-Controlador (MVC)**, que garantiza una separación clara de responsabilidades entre la lógica de IA (modelo), la interfaz de usuario (vista) y la coordinación del flujo (controlador). El producto final será una aplicación web modular, funcional, capaz de detectar contenido falso. Para validar su funcionamiento se realizarán pruebas funcionales, de integración y de precisión del modelo, asegurando la calidad y efectividad del sistema en distintos escenarios.

Palabras clave: Noticias Falsas, Inteligencia Artificial, Procesamiento de Lenguaje Natural (NLP), Aplicación Web



2. Abstract

This project aims to develop a web application with artificial intelligence (AI) based on natural language processing (NLP) to detect fake news. Its importance lies in combating disinformation circulating on social media, promoting access to truthful information, and strengthening trust in digital media. News will be obtained from public online sources, such as news portals, social media, and verified news databases, to feed and train the detection models. The agile Scrum methodology will be used for its development, which allows for flexible, collaborative, and efficient management. Key functionalities such as text analysis and automatic classification will be defined in the Sprint Planning phase. During the Development phase, these tasks will be implemented in short cycles. During the Daily Scrum, the team will review progress and resolve blockers. At the end of each cycle, a Sprint Review will be held to evaluate the results and a Retrospective will be held to propose continuous improvements.

The system will be designed using a layered architecture, organizing the solution into presentation, business logic, and data access layers, which favors scalability and maintenance. The Model-View-Controller (MVC) design pattern will also be applied, ensuring a clear separation of responsibilities between the AI logic (model), the user interface (view), and workflow coordination (controller). The final product will be a modular, functional web application capable of detecting fake content. Functional, integration, and model accuracy tests will be performed to validate its operation, ensuring the system's quality and effectiveness in different scenarios.

Keywords: Fake News, Artificial Intelligence, Natural Language Processing (NLP), Web Application



3. Introducción

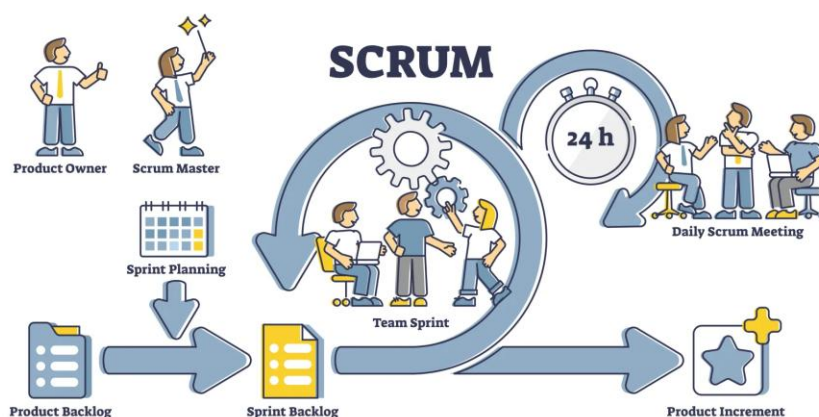
Durante el desarrollo de la aplicación web con inteligencia artificial (IA) basada en procesamiento de lenguaje natural (NPL) para detectar noticias falsas se busca brindar una forma de combatir la desinformación en las redes sociales, permitiendo a los usuarios tener una herramienta que les permita verificar las noticias que leen día a día en sus Facebook o cuentas de X por nombrar algunas plataformas. Al ser una herramienta gratuita podrá ser usada por reporteros o investigadores para verificar fuentes e identificar la confiabilidad de las noticias. Por medio de investigaciones se identificó un proyecto publicado el 31 de mayo del 2023 el cual tiene como título “A Multi-tier Model and Filtering Approach to Detect Fake News Using Machine Learning Algorithms” el cual nos relata como por medio de un modelo multinivel logran etiquetar noticias de acuerdo a ciertos parámetros que ayudan a verificar la veracidad de la información, después de manejar la información por cada capa del modelo, finalmente enviando los resultados a la herramienta WEKA para aprobar los resultados a partir de ambos conjuntos de datos. (Chiung Chang Yu a, 2023)

Teniendo presente la importancia de la información para realizarse su respectivo análisis, se optó por utilizar las redes sociales como Facebook y X (anteriormente llamado Twitter), siendo estos los sitios en donde la información de las noticias suele ser alterada para llamar a la gente y difundir de manera incorrecta los mensajes, la idea de utilizar estas dos redes sociales fue gracias al artículo publicado el 4 de octubre del 2020 el cual lleva como título “Fake News and Aggregated Credibility: Conceptualizing a Co-Creative Medium for Evaluation of Sources Online” en el cual se nos explica como para debatir sobre la credibilidad de las fuentes en línea no es solo suficiente utilizar algoritmos informáticos por lo tanto los autores a lo largo del artículo plantean una propuesta conceptual que busca respaldar la evaluación de la credibilidad de las fuentes en línea utilizando el crowdsourcing (Agnieszka Jaff, 2020).



Para poder llevar a cabo el objetivo principal de este proyecto se eligió utilizar la metodología SCRUM por sus múltiples iteraciones de sus fases que encontramos en la Ilustración 1, las cuales son: **Planificación del Sprint, Desarrollo, Daily Scrum, Revisión del Sprint, retrospectiva.** Gracias a naturaleza de SCRUM que consiste en iteraciones de cuantos Sprint se hubiesen definido en el principio del proyecto permite enfocarse en una parte del proyecto a la vez, dejando así paso para realizar todo por capas de manera efectiva, en caso de no lograr una parte del proyecto en el tiempo esperado, en el Daily Scrum se reajustaran los tiempos para no afectar el tiempo general del proyecto permitiendo así terminar lo que no se logro hacer un su tiempo determinado.

Ilustración 1 Fases de la Metodología SCRUM



Nota: Adaptado de Figura 1, por ILUNION, 2025, <https://www.ilunion.com/es/blog-puntoilunion/scrum>

De la mano de la metodología se implementará la arquitectura por capaz ayudando a orientar los sprints y las tareas planteadas para cumplir los objetivos específicos, esta arquitectura permite la concentración en 3 componentes principales que se seleccionan gracias a un al patrón de diseño seleccionado, siendo este el MVC (Modelo – Vista – Controlador) permitiendo el desarrollo desde las 3 perspectivas necesarias para realizar un proyecto escalable y fácil de mantener en el tiempo. Siguiendo



una definición formal de la arquitectura por capaz, tenemos la siguiente: “Se centra en una distribución jerárquica de las roles y responsabilidades proporcionando una separación efectiva de las preocupaciones (cada cual se encarga de lo que le corresponde)” (costanzo, 2025)

El proyecto se divide en 3 capítulos principales los cuales son la parte del Front-End, Back-End y el modelo encargado de la clasificación de las noticias, cada capítulo tiene sus respectivas actividades para lograr cumplir con el desarrollo de cada sección de la aplicación logrando así cumplir con éxito el objetivo principal del proyecto que es desarrollar una aplicación web con IA basada en NLP utilizando la metodología Scrum para detectar noticias falsas con alta precisión de manera ágil y efectiva.

4. Pregunta de Investigación

En estos días las redes sociales cada vez generan más impacto en las personas, cada vez hay más personas conectadas a la red consumiendo sus servicios entre estos los más destacados son las redes sociales como Facebook y X anteriormente llamada Twitter, plataformas de videos como TikTok y Youtube, siendo estos ejemplos algunos de los conocidos. En estas plataformas ronda mucha información siendo estas plataformas utilizadas para dar propaganda de lugares, comidas o productos, también son utilizadas por medio de comunicación para ayudar a difundir las noticias del día que suceden en el mundo o en las ciudades del país. Estas noticias en su gran mayoría tienen un gran impacto en la audiencia, por esta razón muchas personas con el ánimo de ayudar a confundir y desinformar a la ciudadanía para lograr ciertos crímenes como guiar a grupos de personas a ciertas arias de la ciudad por medio de entrega de ayudas humanitarias falsos o cualquier situación para crear caos.

En la actualidad las redes sociales no ejercen un control como tal sobre la información que las personas publican en las redes sociales, perfectamente una persona puede recrear el formato de publicación de documentos de un periódico y publicarlo como si fuera real, este empezar a ser compartido por las personas que no tiene forma de comprobar su veracidad y muchas veces siendo totalmente

“Formando líderes para la construcción de un nuevo país en paz”



realistas respecto a la sociedad actual, no nos detenemos a buscar en otras fuentes si la información es verídica, simplemente leemos los titulares y damos compartir, ayudan de forma no intencionada con la información. En cada país existen las leyes que castigan a los falsificadores de las plantillas, pero esto en caso de dar con el culpable de la creación de la réplica usada para viralizar las noticias falsas, pero el castigo aplica al simple echo del material registrado por el editorial mas no por la repercusión de lo escrito en las copias. Continuando con la idea de lo limitadas que son las redes sociales en este control de información tomamos como base el siguiente artículo que lleva como título “Language-Independent Fake News Detection: English, Portuguese, and Spanish Mutual Features” en este se nos muestra como por medio de la prosa de cada lenguaje se puede determinar la veracidad de las noticias, en el estudio utilizaron 4 algoritmos de aprendizaje automático para realizar comparación y obtener resultados (HugoQueiroz Abonizio, 2020)

Analizando todo lo descrito anteriormente, lo esperado sería que las redes sociales aplicaran un control sobre toda información catalogada como noticia, pero esto sería un atropello contra la idea principal de las redes sociales en las cuales se supone que son de libre expresión, donde cualquier persona es libre de subir cosas, sea o no información verídica. Para solucionar esto se desearía contar con una herramienta que permita comparar lo descrito en los artículos de noticias con otras fuentes oficiales para establecer si todo lo comunicado en la noticia es verídico o contiene alteraciones para alamar a los lectores. En el artículo que lleva como título “Thefakenewsgraphanalyzer: An open source software for characterizing spreaders in large diffusion graphs” se nos platica de una herramienta de analizar grafos de noticias falsas (FNGA) siendo este un software de código abierto que permite el análisis de extensos grafos de gran tamaño. Esta herramienta ayuda a la investigación de las noticias falsas permitiendo obtener una estadística de cuantas noticias falsas circulan un grupo determinado de información (Amirhosein Bodaghi, 2021).



En la actualidad existen diversos programas que ayudan mayormente en la plataforma de X (Twitter) a distinguir si la cuenta es operada por un Bot en base a los tuits o retuits que realiza la cuenta, en una mejor situación debería existir una página que permita a los usuarios pegar los enlaces de las noticias y analizar la información de la noticia para dar un veredicto en base a unos parámetros de comparación que permita por medio de porcentaje verificar la veracidad de la noticia, permitiendo a los usuarios establecer un veredicto antes de compartir la noticia sin poder verificar la información. Esta aplicación debería tener acceso desde la misma red social para evitar trabajos tediosos como copiar links y cambiar de apps, permitiendo más velocidad a la hora de comprobación.

En base a todo lo planteado surge el problema que la aplicación descrita anteriormente no se ha visto en la actualidad, o por lo menos no ha logrado alcanzar un fuerte auge en el mundo digital, lo que nos sugiere y nos da la oportunidad de preguntarnos lo siguiente: ¿Cómo podemos detectar de manera efectiva cuando una noticia en una red social es totalmente real y su información es completamente veraz?, en base a la incógnita planteada se pensó en el desarrollo de una aplicación web con IA basada en NLP utilizando la metodología Scrum para detectar noticias falsas con alta precisión de manera ágil y efectiva permitiendo a los usuarios combatir la desinformación masiva creada por las redes sociales y la falta de herramientas para evitarla.



5. Justificación

El crear una aplicación web con una IA que usa procesamiento de lenguaje natural (NLP) usando la metodología de Scrum para detectar noticias falsas con alta efectividad es útil desde muchos puntos de vista. Se ve su papel en la sociedad, la manera en que se hace, su aplicación real y cómo ayuda a un desarrollo sustentable según la UNESCO.

En los tiempos de hoy, las falsas noticias se están extendiendo cada vez más y afectando lo que piensa la gente, la paz social y cómo tomamos decisiones sabias. Las noticias falsas causan peleas, afectan las elecciones y hacen peor la confianza en los medios. Esta aplicación ayudará mucho a mejorar la situación prometiendo acceso a verdadera información en redes sociales. Su efecto en la sociedad es ayudar con educación sobre medios y hacer mejor los textos a los cuales el pueblo tiene acceso.

La razón de aplicar la metodología Scrum para el desarrollo del proyecto es que permite un modo iterativo y paso a paso, asegurando la flexibilidad y mejora constante de la aplicación. Scrum ayuda a manejar mejor el proyecto al cortar su desarrollo en sprints, dando una revisión y cambio fijo del producto. Este método apoya la entrega de un sistema que funciona desde inicios, mejorando el tiempo y los recursos usados para tener una solución buena y confiable.

Una razón para la cual la aplicación es viable es desde el punto de vista práctico. Sería una herramienta útil para periodistas, verificadores de información, académicos y usuarios normales. La IA y la NLP permitirán analizar datos a gran escala y establecer patrones en las noticias si son verdaderas y falsas. También será fácil de usar y accesible y podría ser adoptado por muchos medios de comunicación, escuelas y plataformas en línea, por lo que puede ser implementado en el mundo real.

Los ODS que se alinean perfectamente todavía incluyen ODS 16 promover sociedades pacíficas, justas e inclusivas. Deshacerse de la desinformación garantiza transparencia y acceso equitativo a la



información honesta, lo que democratiza los procesos y permite a las personas y organizaciones tomar decisiones informadas. También mejora ODS 4 calidad de la educación. A través de nuestro ejemplo, la sociedad se vuelve más alfabetizada en términos de habilidades de alfabetización e información digital, capacitando a los usuarios para discernir entre información confiable y engañosa. (UNESCO, 2025)

En conclusión, este proyecto representa una solución innovadora y necesaria para combatir la desinformación en la sociedad actual. Su desarrollo mediante Scrum garantiza eficiencia en su implementación, su aplicabilidad práctica lo hace útil para diversos sectores, y su impacto social y sostenible lo convierte en una contribución significativa para la construcción de una sociedad mejor informada y resiliente frente a la manipulación informativa.



6. Objetivos

Objetivo General

Desarrollar aplicación web con IA basada en NLP utilizando la metodología Scrum para detectar noticias falsas con alta precisión de manera ágil y efectiva.

Objetivos Específicos

- Investigar herramientas y tecnologías necesarios para el desarrollo de la aplicación web con IA basada en NLP.
- Establecer las herramientas y tecnologías más precisas para en el desarrollo de la aplicación web con IA basada em NLP.
- Crear la aplicación web con IA basada en NLP.
- Validar la calidad para detectar Fake News de la aplicación web con IA basada en NLP.



7. Estado del arte

7.1 Marco Referencial

7.1.1 Antecedentes Internacionales

- En esta sección contamos con el artículo titulado “Anintelligent cyber threat detection: A swarm-optimized machine learning approach” en el cual se nos propone un sistema de ciberseguridad que se encarga de detectar amenazas, todo esto basado en un aprendizaje automático de enjambre. En este artículo contemplan las Fake News como una amenaza a los sistemas informáticos por la información falsa que proporcionan arruinando sistemas encargados del análisis de datos y demás variedad de softwares. (Issa Qiqieh, 2025)
- Como una referencia importante tenemos el artículo publicado bajo el título “Fake news detection algorithms – A systematic literature review” en el cual se nos presenta una investigación realizada alrededor del tema de las Fake News, buscando en las bibliotecas de Scopus y Science Direct, artículos relacionados con modelos encargados de identificar Fake News, en el artículo se nos presentan las redes neuronales como las predominantes en la detección de las noticias falsas en las plataformas de Facebook y X (Twitter). (Ana Julia Dal Forno, 2025)

7.1.2 Antecedentes Nacionales

- En las investigaciones nacionales tenemos diversas publicaciones como es el caso de artículo titulado “No diga fake news, di desinformación: una revisión sobre el fenómeno de las noticias falsas y sus implicaciones”, en este se nos explica por qué el termino Fake News debería dejar de utilizarse debido a que di desinformación es mas preciso, siendo un término que abarca todas las consecuencias que traen las noticias falsas, en este



artículo se nos deja en claro que esto ya es un problema a nivel mundial que nos afecta a todos. (Pérez, 2019)

- ¿Qué tan involucrados estamos con las noticias falsas?, es una pregunta que nos podemos estar haciendo al momento de pensar en el gran problema que tenemos hoy en día. Un artículo publicado bajo el título “Veracidad y consumo de fake news entre estudiantes de una universidad pública de Cartagena de Indias, Colombia” podemos encontrar mucha incógnitas como la planteada anteriormente, en este artículo se realizaron encuestas aplicadas a los estudiantes de la facultad de Ciencias Económicas de la universidad pública de la ciudad de Cartagena de Indias (Colombia), en donde los resultados obtenidos demostrar como la gran mayoría de la población a esta en contacto con noticias falsas, en varias ocasiones sin saberlo y en otros cuantos casos con el conocimiento de ante mano de falsedad de la información. (Ariel José Torres Pineda, 2021)

7.1.3 Antecedentes Regionales

- A nivel regional también se han publicado algunos artículos relacionados a este tema de la noticias falsas, en esta categoría nos encontramos con un artículo titulado “Análisis de fake news del periódico El Espectador y la Revista Semana sobre el covid-19, y el impacto que generan en los estudiantes de últimos semestres de Comunicación Social de la Universidad de Pamplona, sede Pamplona y la población sanalbertense”, en este artículo se nos explica como en los inicios de la pandemia COVID-19 del año 2020, las noticias falsas respecto a la pandemia eran el pan de cada día, creando caos y una desinformación enorme en la población, en este articulo podemos tener una ejemplo preciso de como las noticias afectan a las comunidades. (Cuadros Bustamante, 2021)



- Al igual que en algunos otros artículos, en el artículo publicado bajo el título “ake news, ¿amenaza u oportunidad para los profesionales de la información y la documentación?”, podemos encontrar una investigación de todos los proyectos y soluciones planteadas hasta la fecha de la publicación del artículo para brindar soluciones a la problemática mundial de las fake news, aprovechando de ser varias entidades las afectadas, se buscaron proyectos que buscaran contrarrestar los efectos de la desinformación. (López-Borrull, 2018)

7.2 Marco Conceptual

Noticias Falsas: Fake News en español significa “Noticias Falsas”. Como su nombre lo indica, son noticias que no son ciertas o que han sido sacadas de contexto. Este es un fenómeno que ha existido desde que el ser humano utiliza el lenguaje para comunicarse. Por ejemplo, los “chismes” generalmente son noticias falsas. Alguien inventó el chisme para su propio beneficio o para perjudicar a otra persona porque se molestó con esa persona. La diferencia es que las “Fake News o Noticias falsas” son chismes o mentiras que se difunden en Internet y en las redes sociales, como por ejemplo en el buscador de Google, Facebook o YouTube. (Rodríguez, 2020)

Inteligencia Artificial: La inteligencia artificial es un campo de la ciencia relacionado con la creación de computadoras y máquinas que pueden razonar, aprender y actuar de una manera que normalmente requeriría inteligencia humana o que involucra datos cuya escala excede lo que los humanos pueden analizar.

La IA es un campo amplio que incluye muchas disciplinas, como la informática, el análisis y la estadística de datos, la ingeniería de hardware y software, la lingüística, la neurociencia y hasta la filosofía y la psicología.



A nivel operativo para el uso empresarial, la IA es un conjunto de tecnologías que se basan principalmente en el aprendizaje automático y el aprendizaje profundo, que se usan para el análisis de datos, la generación de predicciones y previsiones, la categorización de objetos, el procesamiento de lenguaje natural, las recomendaciones, la recuperación inteligente de datos y mucho más. (Google, 2025)

Procesamiento de Lenguaje Natural (NLP): El procesamiento de lenguaje natural (NLP) es una tecnología de machine learning que brinda a las computadoras la capacidad de interpretar, manipular y comprender el lenguaje humano. Hoy en día, las organizaciones tienen grandes volúmenes de datos de voz y texto de varios canales de comunicación, como correos electrónicos, mensajes de texto, fuentes de noticias en redes sociales, vídeo, audio y más. Utilizan software de NLP para procesar de forma automática estos datos, analizan la intención o el sentimiento del mensaje y responden en tiempo real a la comunicación humana. (AWS, 2025)

Aplicación Web: Una aplicación web es un software que se ejecuta en el navegador web. Las empresas tienen que intercambiar información y proporcionar servicios de forma remota. Utilizan aplicaciones web para comunicarse con los clientes cuando lo necesiten y de una forma segura. Las funciones más comunes de los sitios web, como los carros de compra, la búsqueda y el filtrado de productos, la mensajería instantánea y los canales de noticias de las redes sociales, tienen el mismo diseño que las aplicaciones web. Le permiten acceder a funcionalidades complejas sin la necesidad de instalar o configurar un software. (AWS, 2025)

7.3 Marco Contextual

7.3.1 Contexto social y comunicacional

En la actualidad, la circulación masiva de información a través de redes sociales y medios digitales ha transformado la forma en que las personas acceden y consumen noticias. Sin embargo, esta evolución ha traído consigo un fenómeno preocupante: la propagación de noticias falsas (fake news).



Estas publicaciones engañosas no solo afectan la percepción pública sobre temas sensibles, sino que también inciden negativamente en procesos democráticos, decisiones sociales y estabilidad comunitaria. En este contexto, es crucial contar con herramientas tecnológicas que permitan identificar, filtrar y mitigar la desinformación en entornos digitales. El desarrollo de soluciones inteligentes orientadas a esta problemática se convierte en una respuesta concreta ante la necesidad de promover el acceso a información verificada y confiable.

7.3.2 Contexto tecnológico y de desarrollo

El avance de la inteligencia artificial (IA) y el procesamiento de lenguaje natural (NLP) ha permitido crear modelos capaces de interpretar, clasificar y extraer significado de grandes volúmenes de texto. Estas tecnologías, integradas en entornos de desarrollo modernos como Flask para el Back-End y React para el Front-End, permiten construir aplicaciones web eficientes, escalables y orientadas a resolver problemas complejos como la detección de noticias falsas. Mediante la integración de modelos preentrenados como BERT o RoBERTa, junto con librerías como Hugging Face Transformers, TensorFlow o spaCy, es posible implementar soluciones de alta precisión. Además, el uso del patrón de diseño MVC y la arquitectura por capas fortalece la mantenibilidad, escalabilidad y organización del sistema durante su evolución.

7.3.3 Contexto académico y metodológico

Desde una perspectiva académica, este proyecto promueve la aplicación práctica de conocimientos multidisciplinarios que incluyen inteligencia artificial, ingeniería de software, arquitectura de sistemas y metodologías ágiles como Scrum. La elección de esta metodología permite estructurar el trabajo de manera iterativa e incremental, facilitando el control del progreso y la incorporación de mejoras constantes mediante ciclos de revisión. A su vez, se propicia un entorno de colaboración y aprendizaje continuo, clave en el desarrollo de proyectos tecnológicos modernos. En este sentido, la iniciativa no solo



aporta una solución funcional a una problemática real, sino que también representa un ejercicio integral de formación profesional, que refuerza las capacidades técnicas y analíticas requeridas en el ámbito del desarrollo de software con enfoque social.

7.4 Marco Legal

Ley 1581 de 2012: Esta ley establece las disposiciones generales para la protección de datos personales en Colombia. Junto con el Decreto 1377 de 2013, regula el tratamiento de información personal por parte de entidades públicas y privadas. Para efectos del proyecto, es fundamental asegurar que cualquier dato recopilado para el análisis de noticias (que son provenientes de redes sociales de los usuarios) cuente con el consentimiento del titular, y que los datos sean tratados bajo los principios de legalidad, finalidad, libertad, veracidad, transparencia, acceso y seguridad.

8. Metodología

Para la coordinación de las tareas y desarrollo de los objetivos se seleccionó la metodología **SCRUM**, siendo esta la encargada de llevar, controlar y organizar los tiempos de ejecución de cada objetivo general, llegando así al cumplimiento del objetivo general.

8.1 Desde un punto de vista de investigación teórico y filosófico

Paradigma Positivista: El paradigma positivista, también conocido como paradigma cuantitativo, empírico-analítico o racionalista, tiene como objetivo principal explicar, predecir y controlar los fenómenos mediante la verificación de teorías y leyes. Este enfoque es especialmente relevante en la investigación en física y matemática, donde se busca identificar las causas reales de los fenómenos y su relación temporal con otros eventos. (Castrillo, 2024)

Ontología: La ontología o metafísica general es la rama de la filosofía que estudia lo que hay, así como las relaciones entre los entes (por ejemplo, la relación entre un universal —como el rojo— y un



particular que lo "tiene" —como una manzana—) o la relación entre un acto (como el que Sócrates bebiera la cicuta) y sus participantes (Sócrates y la cicuta). (WIKIPEDIA, 2025)

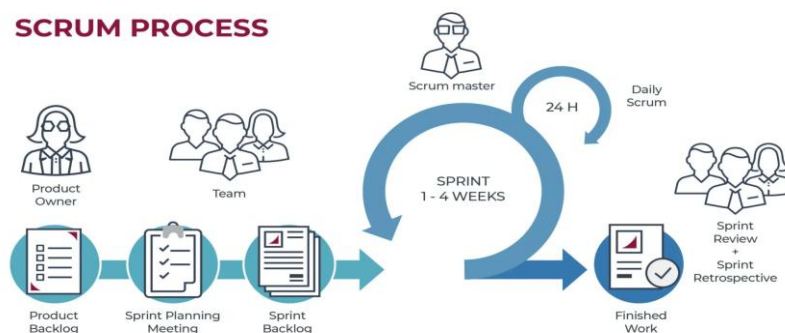
En cuanto al diseño de investigación a utilizar para el desarrollo del proyecto, será el **Descriptivo**, esto por la característica de buscar el detalle en una población o fenómeno, ya sea por medio de censos, estudios prevalencia o cualquier herramienta que permita contar los resultados para tener una estadística de los resultados y poder llegar a una serie de conclusiones.

8.2 Desde el punto de vista de desarrollo

La metodología por medio de sus fases se fusiona con el desarrollo por medio los objetivos específicos, esta relación se expresa de la siguiente manera:

En la ilustración # tenemos las fases de la metodología explicada por partes:

Ilustración 2 Fases de la metodología SCRUM



Nota: Adaptado de ilustración 2, por INITIUM SOFTWARE, 2024,

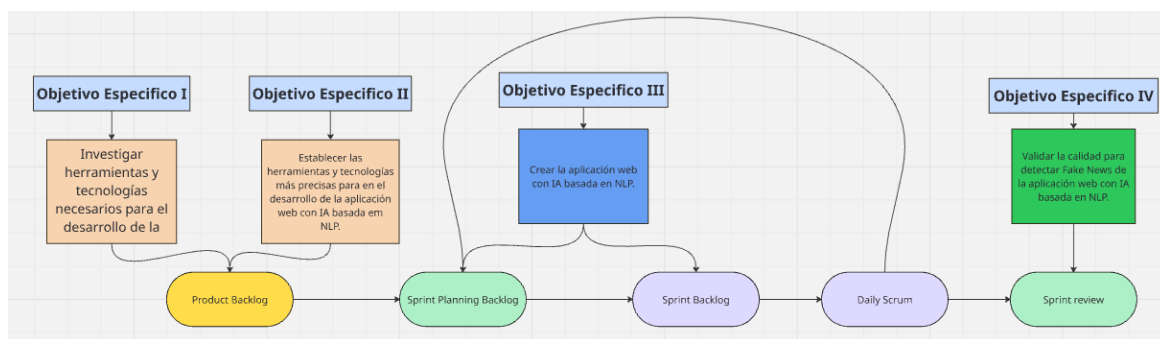
https://www.initiumsoft.com/blog_initium/metodologia-scrum/

En el proyecto contamos con 4 objetivos, centrados en investigar, establecer, crear y validar. Relacionando esto con las fases nos quedarían los primeros dos objetivos que son investigar y establecer, que están directamente relacionados a la fase inicial de la metodología “Product Backlog”, siendo esta



fase en donde se planea y establecen las tecnologías y herramientas para el desarrollo de la aplicación por ende entran las investigaciones previas y la selecciona de todo lo necesario para el desarrollo. La segunda fase “Sprint Planning Backlog” es la encargada de la planificación de los Sprint que se van a ejecutar durante el tiempo del proyecto, esta fase junto a la fase de “Sprint Backlog” que se encarga de ejecutar las tareas planificadas, están relacionadas al objetivo específico 3 que es el encargado del desarrollo de la aplicación web. El objetivo específico 4 siendo de validación esta unido a la fase final de la metodología que es la “Sprint review” en el cual se prueba el resultado y se determina si se realizan cambios y es aprobado el resultado.

Ilustración 3 Relación objetivos con las fases de la metodología



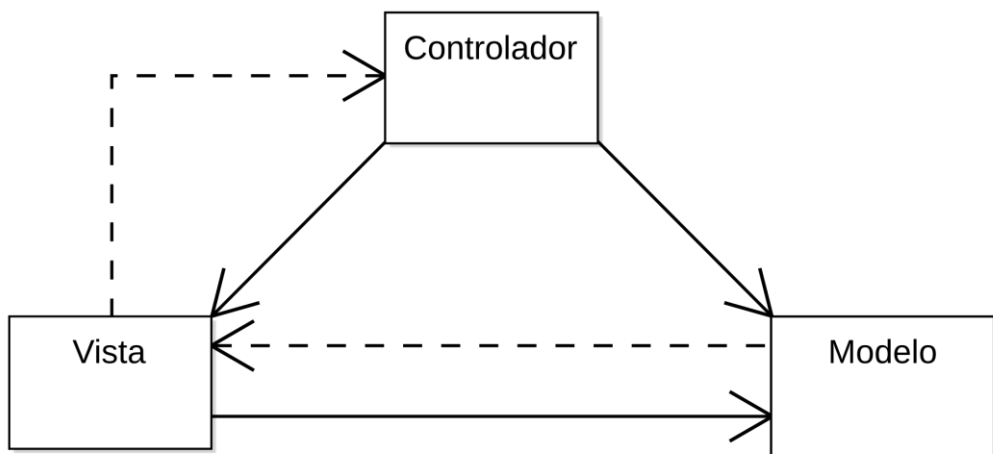
9. Selección del patrón de diseño MVC

El desarrollo de una aplicación web de inteligencia artificial (IA) con procesamiento del lenguaje natural (PLN) para la detección de noticias falsas requiere una arquitectura que garantice la escalabilidad, la facilidad de mantenimiento y una adecuada separación de responsabilidades. Para ello, se ha seleccionado el patrón de diseño Modelo-Vista-Controlador (MVC) como base del sistema. El MVC es un patrón de arquitectura de software, que separa los datos y principalmente lo que es la lógica de negocio de una aplicación de su representación y el módulo encargado de gestionar los eventos y las comunicaciones. Para ello MVC propone la construcción de tres componentes distintos que son el



modelo, la vista y el controlador; es decir: por un lado, define componentes para la representación de la información y, por otro lado, para la interacción del usuario. (Wikipedia, 2025)

Ilustración 4 Modelo Vista Controlador (MCV)



Nota: Esta ilustración muestra cómo se comparte el patrón de diseño MVC

A continuación, se exponen los fundamentos clave de esta decisión.

9.1 Separación de responsabilidades

El patrón MVC permite dividir la aplicación en tres componentes principales:

Modelo: Encargado de la lógica de negocio y la gestión de datos, incluyendo el procesamiento de noticias mediante IA y NLP.

Vista: Responsable de la interfaz de usuario, mostrando los resultados del análisis y proporcionando una experiencia interactiva.

Controlador: Gestiona las solicitudes del usuario, coordinando la comunicación entre la vista y el modelo.



Esta separación facilita el desarrollo modular y la capacidad de realizar cambios en cada componente sin afectar los demás.

9.2 Escalabilidad y mantenimiento

El patrón MVC garantiza la escalabilidad y el mantenimiento de la aplicación, permitiendo la incorporación de nuevas funcionalidades sin afectar toda la estructura del sistema. Si se necesita integrar un nuevo modelo de IA o mejorar la interfaz de usuario, estos cambios pueden realizarse de manera independiente en sus respectivas capas, asegurando un desarrollo modular y flexible. Esta separación facilita la actualización del software y reduce el riesgo de errores al modificar el código. Además, al mantener una organización clara, es más sencillo identificar y corregir fallos, optimizando el rendimiento y asegurando la estabilidad del sistema a largo plazo. (OpenIA, 2025)

9.3 Fácil integración con frameworks y tecnologías modernas

El patrón MVC facilita la integración con frameworks modernos como **Spring Boot (Java)**, **Django (Python)**, **Angular y React (JavaScript)**, permitiendo compatibilidad con herramientas ampliamente utilizadas en el desarrollo web. Esta flexibilidad es esencial para un proyecto que incorpora **inteligencia artificial (IA) y procesamiento de lenguaje natural (PLN)**, ya que posibilita la integración con **TensorFlow, spaCy o Transformers** sin complicaciones. Al estructurar la aplicación en capas, el modelo de IA se gestiona eficientemente en la capa del **modelo**, evitando interferencias con la interfaz de usuario y el control de flujo. Esto garantiza una implementación clara y modular, facilitando futuras actualizaciones y optimizaciones sin afectar la estabilidad del sistema. Además, la compatibilidad con estos frameworks permite aprovechar bibliotecas avanzadas para mejorar el rendimiento y la precisión del análisis de noticias falsas. (OpenIA, 2025)



9.4 Experiencia de usuario mejorada

Al aplicar el patrón MVC, la separación entre la lógica de negocio y la interfaz de usuario permite mejorar la experiencia del usuario sin comprometer la funcionalidad central del sistema. Esto facilita la optimización de la presentación de resultados, asegurando que la información analizada sea clara y comprensible. Además, la interacción con los usuarios se vuelve más dinámica y eficiente, ya que la interfaz puede actualizarse sin afectar el procesamiento de datos en la capa del modelo. Gracias a esta organización, la aplicación puede ofrecer una respuesta rápida y precisa en la detección de noticias falsas, mejorando la usabilidad y garantizando una navegación intuitiva. Esta estructura modular también permite futuras mejoras en la interfaz sin alterar la lógica de negocio, asegurando un desarrollo ágil y adaptable a nuevas necesidades. (OpenIA, 2025)

9.5 Conclusión

El patrón MVC es la mejor opción para el desarrollo de la aplicación debido a su estructura modular, su facilidad de escalabilidad y mantenimiento, su compatibilidad con tecnologías modernas y su capacidad de mejorar la experiencia del usuario. Gracias a esta elección, el sistema podrá evolucionar de manera eficiente, asegurando su efectividad en la detección de noticias falsas y su adaptabilidad a nuevas necesidades en el futuro.

10. Desarrollo

10.1 Objetivo específico I

Para dar solución a la investigación de herramientas y tecnologías necesarios para el desarrollo de la aplicación web con IA basada en NLP se optó por clasificar las tecnologías dependiendo de las capas del desarrollo, estas son en parte Front-End, Back-End y el modelo de clasificación de las noticias. Para comenzar con la investigación se empezó por las herramientas necesarias, como el IDE necesario para el proyecto.



10.1.1 Herramientas

En el aparatado de repositorios tenemos varias opciones, como:

GitHub: Es una forja (plataforma de desarrollo colaborativo) para alojar proyectos utilizando el sistema de control de versiones Git. Se utiliza principalmente para la creación de código fuente de programas de ordenador. (Wikipedia, 2025)

GitLab: ES un servicio web de forja, control de versiones y DevOps basado en Git. Además de gestor de repositorios, el servicio ofrece también alojamiento de wikis y un sistema de seguimiento de errores, todo ello publicado bajo una licencia de código abierto, principalmente. (Wikipedia, 2025)

Bitbucket: Es un servicio de alojamiento basado en web, para los proyectos que utilizan el sistema de control de versiones Mercurial y Git. Bitbucket ofrece planes comerciales y gratuitos. Se ofrece cuentas gratuitas con un número ilimitado de repositorios privados (que puede tener hasta cinco usuarios en el caso de cuentas gratuitas) desde septiembre de 2010, los repositorios privados no se muestran en las páginas de perfil - si un usuario sólo tiene depósitos privados, el sitio web dará el mensaje "Este usuario no tiene repositorios". El servicio está escrito en Python. (Wikipedia, 2024)

En la parte de IDEs tenemos las siguientes opciones:

Visual Studio Code: (también llamado VS Code) es un editor de código fuente desarrollado por Microsoft para Windows, Linux, macOS y Web. Incluye soporte para la depuración, control integrado de Git, resaltado de sintaxis, finalización inteligente de código, fragmentos y refactorización de código. También es personalizable, por lo que los usuarios pueden cambiar el tema del editor, los atajos de teclado y las preferencias. Es software gratuito y de código abierto, aunque la descarga oficial está bajo software privativo e incluye características personalizadas por Microsoft. (Wikipedia, 2025)



IntelliJ IDEA: Es un entorno de desarrollo integrado (IDE) para el desarrollo de programas informáticos. Es desarrollado por JetBrains (anteriormente conocido como IntelliJ), y está disponible en dos ediciones: edición para la comunidad y edición comercial. (Wikipedia, 224)

PyCharm: Es un entorno de desarrollo integrado (IDE) utilizado en programación informática, concretamente para el lenguaje de programación Python. Está desarrollado por la empresa checa JetBrains (antes conocida como IntelliJ). Proporciona análisis de código, un depurador gráfico, un probador de unidades integrado, integración con sistemas de control de versiones (VCS), y soporta el desarrollo web con Django, así como la ciencia de datos con Anaconda. (Wikipedia, 2025)

Para la prueba de conexión con la API tenemos **Postman**: Es una herramienta de colaboración y desarrollo que permite a los desarrolladores interactuar y probar el funcionamiento de servicios web y aplicaciones. proporcionando una interfaz gráfica intuitiva y fácil de usar para enviar solicitudes a servidores web y recibir las respuestas correspondientes. Con esta plataforma se puede gestionar diferentes entornos de desarrollo, organizar las solicitudes en colecciones y realizar pruebas automatizadas para verificar el comportamiento de los sistemas. **Postman** es utilizado por los desarrolladores para testear colecciones y catálogos APIs (tanto a nivel front-end como back-end), para gestionar el ciclo de vida de las APIs, mejorar el trabajo colaborativo y mejorar la organización del proceso de diseño y desarrollo. (Javier, 2023)

Para la gestión de las tareas del equipo por medio de la metodología Scrum tenemos los siguientes softwares:

Trello: Es una herramienta de gestión de proyectos y tareas basada en la nube que utiliza el sistema Kanban para facilitar la organización y colaboración visual de equipos. Permite visualizar el flujo



de trabajo mediante tableros, listas y tarjetas, lo que facilita la gestión de tareas y el seguimiento del progreso. (google, 2025)

Jira: Es un producto de software propietario para la gestión de proyectos, seguimiento de errores e incidencias. La herramienta fue desarrollada por la empresa australiana Atlassian. Inicialmente se utilizó para el desarrollo de software, sirviendo de apoyo para la gestión de requisitos, seguimiento del estado de desarrollo y más tarde para la gestión de errores. Jira puede ser utilizado para la gestión y mejora de los procesos, gracias a sus funciones para la organización de flujos de trabajo. (Wikipedia, 2024)

Para la representación del modelo de usuario de la aplicación tenemos **Figma:** Es un editor de gráficos vectorial y una herramienta de generación de prototipos, principalmente basada en la web, con características off-line adicionales habilitadas por aplicaciones de escritorio en macOS y Windows. Las aplicaciones Figma Mirror companion para Android y iOS permiten mirar los prototipos de Figma en dispositivos móviles. El conjunto de funciones de Figma, se enfoca en el uso de la interfaz de usuario y el diseño de experiencia de usuario, con énfasis en la colaboración en tiempo real. (Wikipedia, 2025)

10.1.2 Tecnologías

Entrando en la parte de las tecnologías a utilizar en el desarrollo de la aplicación tenemos las opciones de lenguajes y estructuras a utilizar en el Front-End:

HTML5: HTML5 es la última versión de HyperText Markup Language (Lenguaje de Marcado de Hipertexto). Combina tres tecnologías web: HTML, para construir la estructura de la página; CSS, para mejorar la presentación; y JavaScript, para agregar funcionalidad a los elementos HTML. Sus etiquetas o marcas hacen que textos, imágenes y elementos multimedia se visualicen correctamente en la pantalla, y que las páginas web tengan un formato unificado, estandarizado y sean compatibles con distintos navegadores, como Chrome, Edge, Firefox y Safari. (EBAC, 2023)



CSS3: (siglas en inglés de Cascading Style Sheets; en español: «hojas de estilo en cascada») es un lenguaje informático especializado en definir y cohesionar la presentación de un documento escrito en un lenguaje de marcado como HTML o XML; por ello, puede ser aplicado a cualquier documento etiquetado en XHTML, SVG, XUL, RSS, etcétera. Junto con el propio HTML y JavaScript, CSS está en la base de la tecnología usada por muchos sitios web para crear páginas visualmente atractivas, interfaces de usuario para aplicaciones web y GUIs para aplicaciones móviles. (Wikipedia, 2025)

Tailwind CSS: Es un framework de CSS de código abierto para el diseño de páginas web. La principal característica de esta biblioteca es que, a diferencia de otras como Bootstrap, no genera una serie de clases predefinidas para elementos como botones o tablas. En su lugar, crea una lista de clases CSS "de utilidad" que se pueden usar para dar estilos individuales a cada elemento. (Wikipedia, 2025)

Bootstrap: Es un framework multiplataforma o conjunto de herramientas de código abierto para diseño de sitios y aplicaciones web. Contiene plantillas de diseño con tipografía, formularios, botones, cuadros, menús de navegación y otros elementos de diseño basado en HTML y CSS, así como extensiones de JavaScript Turner Broadcasting System EMEA adicionales. A diferencia de muchos frameworks web, solo se ocupa del desarrollo front-end. (Wikipedia, 2025)

JavaScript: (abreviado comúnmente JS) es un lenguaje de programación interpretado, dialecto del estándar ECMAScript. Se define como orientado a objetos, basado en prototipos, imperativo, débilmente tipado y dinámico. Se utiliza principalmente del lado del cliente, implementado como parte de un navegador web permitiendo mejoras en la interfaz de usuario y páginas web dinámicas y JavaScript del lado del servidor (Server-side JavaScript o SSJS). Su uso en aplicaciones externas a la web, por ejemplo, en documentos PDF, aplicaciones de escritorio (mayoritariamente widgets) es también significativo. (Wikipedia, 2025)



React: (también llamada React.js o ReactJS) es una biblioteca Javascript de código abierto diseñada para crear interfaces de usuario con el objetivo de facilitar el desarrollo de aplicaciones en una sola página. React intenta ayudar a los desarrolladores a construir aplicaciones que usan datos que cambian todo el tiempo. Su objetivo es ser sencillo, declarativo y fácil de combinar. React sólo maneja la interfaz de usuario en una aplicación; React es la Vista en un contexto en el que se use el patrón MVC (Modelo-Vista-Controlador) o MVVM (Modelo-vista-modelo de vista). También puede ser utilizado con las extensiones de React-based que se encargan de las partes no-UI (que no forman parte de la interfaz de usuario) de una aplicación web. (Wikipedia, 2024)

Angular: (comúnmente llamado Angular 2+ o Angular 2) es un framework para aplicaciones web desarrollado en TypeScript, de código abierto, mantenido por Google, que se utiliza para crear y mantener aplicaciones web de una sola página. Su objetivo es aumentar las aplicaciones basadas en navegador con capacidad de Modelo Vista Controlador (MVC), en un esfuerzo para hacer que el desarrollo y las pruebas sean más fáciles. (Wikipedia, 2024)

Vue.js: (comúnmente conocido como Vue) es un framework de JavaScript de código abierto para la construcción de interfaces de usuario y aplicaciones de una sola página. Vue.js cuenta con una arquitectura de adaptación gradual que se centra en la representación declarativa y la composición de componentes. La biblioteca central se centra sólo en la capa de vista. Las características avanzadas necesarias para aplicaciones complejas como el enrutamiento, la gestión de estados y las herramientas de construcción se ofrecen a través de librerías y paquetes de apoyo mantenidos oficialmente con Next.js como una de las soluciones más populares. (Wikipedia, 2025)

Continuando con la investigación de tecnologías para el desarrollo debemos visualizar las opciones para el desarrollo de la parte Back-End del proyecto, entre estas opciones encontramos:



Node.js: Es un entorno de ejecución de un solo hilo, de código abierto y multiplataforma para crear aplicaciones de red y del lado del servidor rápidas y escalables. Se ejecuta en el motor de ejecución de JavaScript V8, y utiliza una arquitectura de E/S basada en eventos y sin bloqueos, lo que la hace eficiente y adecuada para aplicaciones en tiempo real. (KINSTA, 2025)

Spring Boot: es una herramienta que hace que el desarrollo de aplicaciones web y microservicios con Spring Framework sea más rápido y fácil a través de tres funcionalidades principales: configuración automática, enfoque obstinado de la configuración, capacidad de crear aplicaciones independientes. Estas características funcionan juntas para brindarle una herramienta que le permite configurar una aplicación basada en Spring con una configuración y preparación mínimas. (IBM, 2025)

Django: es un framework de desarrollo web de código abierto, escrito en Python, que respeta el patrón de diseño conocido como modelo–vista–controlador (MVC). La meta fundamental de Django es facilitar la creación de sitios web complejos. Django pone énfasis en el re-uso, la conectividad y extensibilidad de componentes, el desarrollo rápido y el principio «DRY» (del inglés Don't Repeat Yourself, «No te repitas»). El lenguaje Python es usado en todos los componentes del framework, incluso en configuraciones, archivos, y en sus modelos de datos. (Wikipedia, 2025)

Flask: Es un marco de desarrollo web de código abierto, escrito en Python. Fue diseñado para ser un marco minimalista, flexible y fácil de usar para la creación de aplicaciones web. Proporciona una manera sencilla de crear y desplegar aplicaciones web dinámicas; permite que los desarrolladores se centren en la lógica de la aplicación en lugar de que se preocupen por la infraestructura subyacente. Asimismo, ofrece una gran cantidad de libertad y control sobre el desarrollo de la aplicación, lo que lo vuelve ideal para proyectos pequeños y medianos. (mytaskpanel, 2025)



Express.js: también conocido como "Express", es un framework de backend de Node.js minimalista, rápido y similar a Sinatra. Ofrece funciones y herramientas robustas para desarrollar aplicaciones escalables. Ofrece un sistema de enrutamiento y funciones simplificadas para ampliar el framework mediante el desarrollo de componentes y partes más potentes según los casos de uso de su aplicación. (KINSTA, 2023)

MongoDB: Es un sistema de gestión de bases de datos (DBMS, por sus siglas en inglés) no relacionales y de código abierto, que utiliza documentos flexibles en lugar de tablas y filas para procesar y almacenar varias formas de datos. Al ser una solución de base de datos NoSQL, MongoDB no requiere un sistema de gestión de bases de datos relacionales (RDBMS), por lo que proporciona un modelo de almacenamiento de datos elástico que permite a los usuarios almacenar y consultar fácilmente tipos de datos multivariados. Esto no solo simplifica la gestión de la base de datos para los desarrolladores, sino que también crea un entorno altamente escalable para aplicaciones y servicios multiplataforma. (IBM, 2025)

PostgreSQL: Es una base de datos de código abierto que tiene una sólida reputación por su fiabilidad, flexibilidad y soporte de estándares técnicos abiertos. A diferencia de otros RDBMS (sistemas de gestión de bases de datos relacionales), PostgreSQL (enlace externo a ibm.com) soporta tipos de datos relacionales y no relacionales. Esto la convierte en una de las bases de datos relacionales más compatibles, estables y maduras disponibles actualmente. (IBM, 2025)

MySQL: Es el sistema de gestión de bases de datos relacional más extendido en la actualidad al estar basada en código abierto. MySQL es un sistema de gestión de bases de datos que cuenta con una doble licencia. Por una parte, es de código abierto, pero por otra, cuenta con una versión comercial gestionada por la compañía Oracle. (Robledo, 2019)



Continuando con una parte importantísima del desarrollo, se presentan las tecnologías para el desarrollo del Modelo de Clasificación de las noticias:

Python: Es un lenguaje de programación ampliamente utilizado en las aplicaciones web, el desarrollo de software, la ciencia de datos y el machine learning (ML). Los desarrolladores utilizan Python porque es eficiente y fácil de aprender, además de que se puede ejecutar en muchas plataformas diferentes. El software Python se puede descargar gratis, se integra bien a todos los tipos de sistemas y aumenta la velocidad del desarrollo. (AWS, 2025)

Scikit-learn: (anteriormente scikits.learn y también conocido como sklearn) es una biblioteca de aprendizaje automático gratuita y de código abierto para el lenguaje de programación Python. Incluye diversos algoritmos de clasificación, regresión y agrupamiento, incluyendo máquinas de vectores de soporte, bosques aleatorios, potenciación de gradiente, k -medias y DBSCAN, y está diseñado para interoperar con las bibliotecas numéricas y científicas de Python NumPy y SciPy. Scikit-learn es un proyecto financiado con fondos públicos por NumFOCUS. (Wikipedia, 2025)

spaCy: Es una librería de software para procesamiento de lenguajes naturales, reconocimiento de nombres de entidades, análisis de redes, visualización de datos, análisis, visual analysis, análisis de contenidos, enriching, anotación desarrollada por Matt Honnibal y programado en lenguaje Python. (Wikipedia, 2022)

NLTK: El kit de herramientas de lenguaje natural, o más comúnmente NLTK, es un conjunto de bibliotecas y programas para el procesamiento del lenguaje natural (PLN) simbólico y estadísticos para el lenguaje de programación Python. NLTK incluye demostraciones gráficas y datos de muestra. Se acompaña de un libro que explica los conceptos subyacentes a las tareas de procesamiento del lenguaje compatibles el toolkit, además de programas de ejemplo. NLTK está destinado a apoyar la investigación y



la enseñanza en procesamiento de lenguaje natural (PLN) o áreas muy relacionadas, que incluyen la lingüística empírica, las ciencias cognitivas, la inteligencia artificial, la recuperación de información, y el aprendizaje de la máquina. NLTK se ha utilizado con éxito como herramienta de enseñanza, como una herramienta de estudio individual, y como plataforma para los sistemas de investigación de prototipos y construcción. (Wikipedia, 2024)

TensorFlow: Es una biblioteca de código abierto para cálculos numéricos, aprendizaje automático a gran escala, el Deep learning y otras actividades de análisis estadístico y predictivo. Esta herramienta facilita la implementación de modelos de machine learning para los desarrolladores, asistiéndoles en el proceso de adquisición de datos, formulación de predicciones a gran escala y el afinamiento de los resultados. (EAE, 2025)

PyTorch: Es un marco de deep learning de código abierto basado en software que se utiliza para crear redes neuronales, combinando la biblioteca de machine learning (ML) de Torch con una API de alto nivel basada en Python. Su flexibilidad y facilidad de uso, entre otros beneficios, lo han convertido en el marco de ML líder para las comunidades académicas y de investigación. (Dave Bergmann, 2023)

Transformers (Hugging Face): Es una biblioteca de modelos preentrenados de procesamiento de lenguaje natural, visión artificial, audio y multimodales para inferencia y entrenamiento. Use Transformers para entrenar modelos con sus datos, crear aplicaciones de inferencia y generar texto con modelos de lenguaje extensos. (Hugging face, 2025)

Pandas: Es una biblioteca de software escrita como extensión de Numpy para manipulación y análisis de datos para el lenguaje de programación Python. En particular, ofrece estructuras de datos y operaciones para manipular tablas numéricas y series temporales. (Universidad de Alcalá, 2025)



NumPy: Es una biblioteca de Python utilizada para trabajar con matrices. También tiene funciones para trabajar en el dominio del álgebra lineal, transformada de Fourier y matrices. NumPy fue creado en 2005 por Travis Oliphant. Es un proyecto de código abierto y se puede usar libremente. (W3School, 2025)

10.2 Objetivo Especifico II

Teniendo en cuenta toda la información consultada con anticipación, se puede tomar una decisión respecto a las herramientas y tecnologías más precisas para cada capa en el desarrollo de la aplicación web con IA basada em NLP.

10.2.1 Selección de las Herramientas

Para seleccionar de las herramientas se realizaron cuadros comparativos de ventajas y desventajas, para lograr establecer el mejor para el desarrollo.

Tabla 1 Comparación de Herramientas de versiones

Herramienta	Ventajas	Desventajas
GitHub	<ul style="list-style-type: none"> - Comunidad más grande y activa - Integración directa con Actions (CI/CD) - Compatible con Hugging Face, VS Code y otras herramientas - Amplia documentación y recursos - Interfaz intuitiva y moderna 	<ul style="list-style-type: none"> - Algunas funciones avanzadas (como seguridad avanzada) requieren GitHub Pro
GitLab	<ul style="list-style-type: none"> - CI/CD integrado más completo en su versión gratuita - Permite repos privados ilimitados - Se puede instalar en servidores propios (self-hosted) 	<ul style="list-style-type: none"> - Interfaz más compleja para nuevos usuarios - Menor comunidad para proyectos de IA/NLP



Bitbucket	<ul style="list-style-type: none"> - Integración nativa con Jira y Trello - Repositorios privados ilimitados en plan gratuito - Adecuado para proyectos pequeños y medianos 	<ul style="list-style-type: none"> - Menor comunidad open-source - Pocas integraciones específicas para proyectos de IA - Interfaz menos intuitiva
------------------	--	---

Nota: En la tabla podemos encontrar las ventajas y desventajas de cada controlador de versiones investigado en el objetivo específico I

En base a la información de la tabla 1 podemos concluir que el controlador de versiones a utilizar en el proyecto será **GitHub**.

Para la selección del IDE se realizó una tabla de comparación del IDE con más compatibilidad con las tecnologías a utilizar en cada capa del desarrollo.

Tabla 2 Comparación de IDEs con tecnologías

IDE	Tecnologías						
	HTML/CSS/JS	React / Angular	Django / Flask	Spring Boot	Python (IA/NLP)	TensorFlow / PyTorch	Integración con Git
Visual Studio Code	Excelente soporte	Muy buena con extensiones	Soporte completo	Requiere configuración adicional	Completo	Con plugins	Nativo y fluido
IntelliJ IDEA	Compatible con plugins	Compatible con plugins	Mejor con plugins adicionales	Soporte nativo (ideal para Java)	Básico (no es su fuerte)	Requiere configuración	Integración nativa
PyCharm	Limitado (más básico)	Menor compatibilidad	Excelente (ideal para Django/Flask)	No recomendado para Java	Ideal para Python	Soporte completo	Muy buen soporte

Nota: En la tabla encontramos la comparación de los IDEs para trabajar con todas las tecnologías consultadas en el desarrollo del Objetivo Especifico I



En base a la información plasmada en la tabla 2 se va a utilizar **Visual Studio Code (VSCode)** para el desarrollo de cada una de las capas del proyecto, esto por su casi completa compatibilidad con las tecnologías investigadas.

La herramienta de **Postman** será utilizada para realizar diversas pruebas de respuesta del servidor una vez conectado con la parte del Front-End, esto para comprobar la efectividad de los métodos y se logren cumplir los resultados esperados.

Para seleccionar la herramienta para llevar el control de las actividades programadas por la metodología SCRUM se realizó una tabla comparativa entre los dos softwares investigados.

Tabla 3 Comparación entre los Software de manejo de SCRUM

Característica	Trello	Jira
Facilidad de uso	Muy fácil e intuitivo	Más complejo, curva de aprendizaje más alta
Soporte para Scrum	Limitado (requiere plantillas o Power-Ups)	Soporte nativo (backlogs, sprints, epics, etc.)
Gestión de tareas	Listas y tarjetas personalizables	Completa y estructurada con flujos de trabajo
Seguimiento de progreso	Básico (tableros Kanban simples)	Avanzado (burndown charts, velocity, etc.)
Integraciones	Amplias (Slack, GitHub, Google Drive, etc.)	Excelente (con Bitbucket, GitHub, Confluence, etc.)
Escalabilidad para equipos grandes	Limitada	Alta, ideal para equipos de desarrollo
Coste	Gratuito con funciones básicas	Gratuito hasta 10 usuarios, luego es de pago
Automatización	Mediante Power-Ups (Butler)	Flujo de trabajo automatizado más robusto
Visualización del proyecto	Intuitiva (tableros tipo Kanban)	Detallada, aunque más técnica



Recomendado para	Equipos pequeños o proyectos simples	Equipos de desarrollo ágil, proyectos complejos
------------------	--------------------------------------	---

Nota: En la tabla se presentan las características principales de cada software de control de la metodología SCRUM.

En base a la información proporcionada en la tabla 3 para el manejo de los tiempos del proyecto se seleccionó **Trello** siendo este el software que más se ajusta a las características del proyecto.

Para la representación de las interfases de usuario se eligió la herramienta de **Figma**, esto por su fácil manejo y posibilidad de mostrar el desplazamiento entre ventanas del desarrollo, logrando darles a los usuarios interesados una perspectiva de cómo será el resultado final.

10.2.2 Selección de las Tecnologías

Como parte del desarrollo del segundo objetivo específico del proyecto presentamos la información para decidir cual tecnología es mas apropiada para capa del desarrollo.

Tecnología para el Front-End

Tabla 4 Fortalezas de cada tecnología de desarrollo

Tecnología	Fortalezas principales
HTML/CSS/JavaScript	<ul style="list-style-type: none"> - Base fundamental del desarrollo web - Compatible con todos los navegadores - Fácil de aprender e implementar - Ideal para prototipos rápidos y proyectos simples
React	<ul style="list-style-type: none"> - Reutilización de componentes - Alto rendimiento con Virtual DOM - Gran ecosistema (Next.js, Redux, etc.) - Flexible e integrable con backends y APIs - Muy popular en aplicaciones modernas
Angular	<ul style="list-style-type: none"> - Arquitectura completa para aplicaciones grandes



	<ul style="list-style-type: none"> - Inyección de dependencias y TypeScript nativos - Vinculación bidireccional de datos - Ideal para aplicaciones empresariales estructuradas
Vue.js	<ul style="list-style-type: none"> - Curva de aprendizaje suave - Sintaxis clara y fácil de usar - Requiere menos configuración inicial que Angular - Componente reactivo similar a React, pero más sencillo - Ideal para proyectos escalables con bajo nivel de complejidad inicial

Nota: En la tabla se nos presentan las fortalezas de cada tecnología para el desarrollo del Front-End del proyecto.

En base la información suministrada en la tabla 4 podemos concluir lo siguiente:

HTML/CSS/JS: Base necesaria y útil en cualquier stack, pero limitado sin un framework.

React: Recomendado para proyectos modernos, modulares y dinámicos como el proyecto.

Angular: Más adecuado para equipos grandes y aplicaciones empresariales estructuradas.

Vue.js: Perfecto si buscas una combinación de simplicidad y escalabilidad.

De acuerdo al análisis realizado, se optó por utilizar **React** para el desarrollo de la parte Front-End del proyecto, esto por su compatibilidad con el objetivo del proyecto.

Tecnología para el Back-End

Tabla 5 Comparación de las fortalezas de cada tecnología para el Back-End

Tecnología	Fortalezas principales
Django	<ul style="list-style-type: none"> - Framework robusto con ORM, autenticación y admin integrado - Excelente seguridad por defecto - Ideal para proyectos con IA en Python



Flask	<ul style="list-style-type: none"> - Microframework ligero y flexible - Ideal para APIs rápidas - Muy usado en proyectos con IA/NLP
Spring Boot	<ul style="list-style-type: none"> - Escalable y estructurado - Ideal para entornos empresariales - Soporte nativo para servicios REST
Node.js	<ul style="list-style-type: none"> - Entorno de ejecución rápido y eficiente basado en JavaScript - Ideal para manejo de múltiples conexiones y eventos - Compatible con todo el ecosistema JS
Express.js	<ul style="list-style-type: none"> - Framework minimalista para Node.js - Permite crear APIs REST de forma rápida y flexible - Gran comunidad y soporte

Nota: En la tabla se nos presenta las fortalezas de cada tecnología investigada para el Bach-End.

En conformidad con la información presentada en la tabla 5 podemos concluir lo siguiente:

Node.js + Express.js ofrecen una integración fluida y nativa gracias al uso de JavaScript.

Django y Flask son potentes para proyectos que integran IA con Python.

Spring Boot es recomendable para aplicaciones robustas con Java.

De acuerdo a la información recolectada la tecnología seleccionada para el desarrollo Back-End de la aplicación va a ser Flask por su flexibilidad y rapidez a la hora de los desarrollos.

Tecnología para el modelo de clasificación de Fake News

Tabla 6 Comparación de Modelos y librerías de Clasificación

Herramienta / Librería	Propósito principal	Fortalezas clave
Python	Lenguaje de programación base	<ul style="list-style-type: none"> - Sintaxis simple y legible - Gran ecosistema de librerías de IA/NLP
Scikit-learn	Machine learning tradicional	<ul style="list-style-type: none"> - Ideal para modelos clásicos (Naive Bayes, SVM)



		- Fácil integración con Pandas
spaCy	Procesamiento de lenguaje natural	- Rápido y eficiente para tareas NLP básicas - Buen soporte para pipelines
NLTK	Procesamiento lingüístico académico	- Extensa base de datos lingüísticos - Ideal para prototipos y aprendizaje
TensorFlow	Framework de deep learning	- Alta eficiencia en redes neuronales - Gran comunidad y herramientas de producción
PyTorch	Framework de deep learning	- Más flexible y dinámico que TensorFlow - Preferido en investigación
Transformers (Hugging Face)	Modelos preentrenados para NLP (BERT, RoBERTa, etc.)	- Alta precisión en clasificación de texto - Compatible con PyTorch y TensorFlow
Pandas	Manipulación de datos tabulares	- Ideal para preprocesar datasets - Integra con Scikit - learn y NumPy
NumPy	Cálculo numérico y matrices	- Base matemática de otras librerías - Rápido y versátil

Nota: En la tabla podemos encontrar la comparación de todos los modelos y librerías de clasificación compatibles con el proyecto.

En base a la información recopilada, se consideró la creación de un modelo personal que evalué ciertos parámetros para dar la clasificación de las noticias, por lo tanto, el modelo será realizado bajo el lenguaje en general de Python con las librerías que se vayan necesitando a lo largo de los cambios en la idea principal del modelo.

10.3 Objetivo Especifico III

Para el cumplimiento del objetivo específico número III el cual consiste en crear la aplicación web con IA basada en NLP, se dividió este objetivo en 3 fases para mejor tiempo de desarrollo y

“Formando líderes para la construcción de un nuevo país en paz”



ejecución en el proyecto. Las fases selectas fueron las ya mencionadas anteriormente en el documento **(Front-End, Back-End y Modelo de clasificación)**. Para el inicio del proyecto se comenzó por la parte Front-End para dar a los usuarios un adelanto de la orientación de los elementos y dar a observar las funcionalidades que se podrán utilizar cuando el desarrollo este culminado.

10.3.1 Front-End

Para el desarrollo Front-End se eligió la tecnología de React, bajo esta se realizó la interfaz gráfica que el usuario la a estar utilizando, esta interfaz la podemos observar la ilustración 5

Ilustración 5 Interfaz Gráfica de Detector Fake News

The screenshot shows a web application titled "Detector Fake News". It has a green header with the title and a magnifying glass icon. Below the header is a white box with the heading "Información" and the text "No desconfíes más de la información de tus redes sociales. Pega el link de la noticia y verifica su veracidad." There is a link input field with the placeholder "Link" and a "Verificar" button. Below this is a green box with the heading "Resultados". It contains a table with the following data:

Notas de la evaluación	resultado
Utiliza palabras clave	SI
Presenta imágenes que corroboran la información	SI
El emisor de la noticia es una página o persona dedicada a compartir noticias	NO

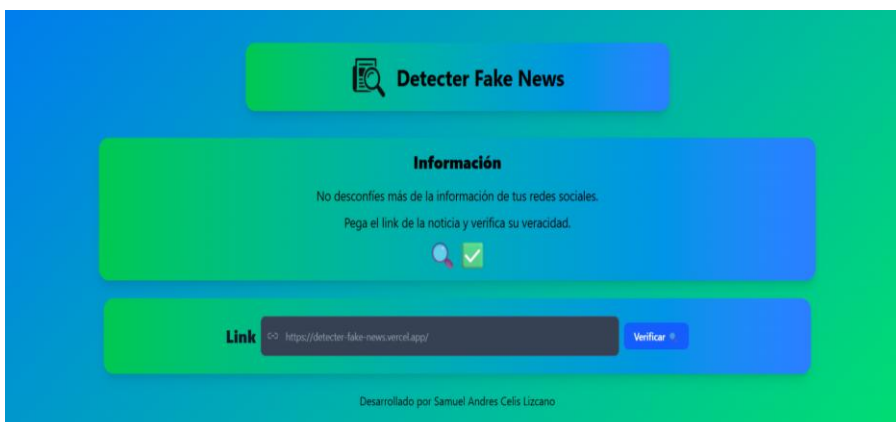
Below the table is a "Conclusión" section with the text: "De acuerdo al análisis realizado y teniendo en cuenta los ítems de evaluación la noticia es verdadera. Recuerda que este herramienta solo es una Bata y no debe usarse como única herramienta en la evaluación de las noticias." At the bottom, it says "Desarrollado por Samuel Andres Calle Lizaro".

En la ilustración 5 podemos visualizar varios elementos, entre estos tenemos un apartado para el título de la aplicación, un espacio para la información o descripción de la aplicación, un espacio para agregar el link de la noticia y un botón para realizar la comprobación y en la parte final tenemos un espacio en donde se nos muestra una tabla con unos parámetros los cuales se aprobarán o negarán dependiendo de los resultados del análisis en el modelo. Cabe resalta que el apartado de Resultados solo



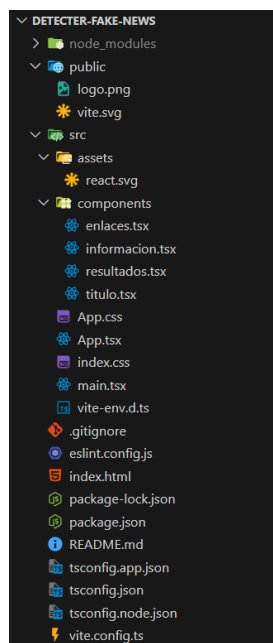
aparece una vez se haya realizado el análisis de una noticia, de forma inicial este campo no aparecerá como se muestra en la ilustración 6.

Ilustración 6 Pagina inicial



Para el caso de la construcción de la parte Front-End el proyecto se estructuro de la forma como lo muestra la ilustración 7.

Ilustración 7 Estructura de carpetas Front-End





Como camino principal tenemos la siguiente secuencia de archivos: index.html se comunica con main.tsx, main.tsx se comunica con app.tsx para finalmente este hacer los respectivos llamados a los componentes que le dan vida y funcionalidades a la aplicación web, en la ilustración 8 podemos observar como app.tsx unifica todos los componentes.

Ilustración 8 Archivo app.tsx

```
src > App.tsx > ...
You, hace 5 días | 1 author (You)
1 // import { useState } from 'react'
2 import './App.css'
3 import Titulo from './components/titulo';
4 import Informacion from './components/informacion';
5 import Enlaces from './components/enlaces';
6 import Resultados from './components/resultados';
7
8 function App() {
9
10   return (
11     <>
12       <Titulo />
13       <Informacion />
14       <Enlaces />
15       <Resultados />
16       <p className="text-black">
17         Desarrollado por Samuel Andres Celis Lizcano
18       </p>
19     </>
20   )
21 }
22
23
24 export default App
25
```

Nota: En la ilustración podemos evidenciar en la primera línea del archivo comentada, esta importación fue habilitada cuando se estableció la conexión con la parte Back-End.

Para comentar un poco de la composición de los componentes de la aplicación, en la ilustración 9 encontraremos como esta creado el componente de Enlaces.



Ilustración 9 Componente Enlaces

```
src > components > enlaces.tsx > ...
You, hace 16 horas | 1 author (You)
1 export default function Enlaces(){
2   return (
3     <div className="m-6 flex flex-col flex-wrap items-center justify-center p-4
4       rounded-2xl bg-gradient-to-l from-blue-500 to-green-500 shadow-lg
5       hover:shadow-2xl hover:shadow-white hover:text-white hover:border sm:flex-row">
6       <p className="text-2xl font-extrabold text-gray-900 justify-center
7         items-center flex">Link</p>
8       <div className="pr-2 max-w-[600px] w-full relative">
9         <div className="absolute inset-y-0 start-0 flex items-center ps-5
10           pointer-events-none">
11           { /* Icono de enlace */ }
12           <svg className="h-4 text-gray-500 dark:text-gray-400"
13             aria-hidden="true" xmlns="http://www.w3.org/2000/svg" fill="none"
14             viewBox="0 0 24 24">
15             <path stroke="currentColor" strokeWidth="2" strokeLinecap="round"
16               strokeLinejoin="round" d="M17 7h2a5 5 0 0 1 0 10h-2m-10 0H5a5 5 0
17               0 1 0-10h2m2 5h6"/>
18           </svg>
19         </div>
20         <input type="search" id="search" className="m-2 block w-full p-4 ps-10
21           text-sm text-gray-900 border border-gray-300 rounded-lg bg-gray-50
22           focus:ring-blue-500 focus:border-blue-500 dark:bg-gray-700
23           dark:border-gray-600 dark:placeholder-gray-400 dark:text-white
24           dark:focus:ring-blue-500 dark:focus:border-blue-500" placeholder="https://
25             detector-fake-news.vercel.app/" required />
26         </div>
27         <button className="text-white bg-blue-700 hover:bg-blue-800 focus:ring-4
28           focus:outline-none focus:ring-blue-300 font-medium rounded-lg text-sm px-4
29           py-2 ml-2 dark:bg-blue-600 dark:hover:bg-blue-700 dark:focus:ring-blue-800
30           cursor-pointer max-w-[160px]">
31           Verificar<span className="animate-pulse"> 🔍 </span>
32         </button>
33       </div>
34     </div>
35   )
36 }
```

Nota: Podemos evidenciar como esta creado el componente de enlaces de la aplicación web, el cual consta de varios elementos div y p, en esta estructura también destacan los elementos input y button encargados de dar la funcionalidad de la aplicación web.

10.3.2 Back-End

Para la parte del desarrollo Back-End, siendo esta parte la segunda fase del desarrollo de la aplicación web, tenemos una conexión muy simple, esto debido a que no se necesita un sistema de base de datos en la aplicación, solo se necesita establecer la conexión con el modelo y ya. Para dar cumplimiento a esto se utiliza una función llamada analizar con el método POST para realizar la debida conexión, en la ilustración 10 podemos observar como esta compuesto el Back-End de la aplicación.



Ilustración 10 Código Back-End

```

1  # backend/app.py
2  from flask import Flask, request, jsonify
3  from flask_cors import CORS
4  from modelo import analizar_noticia
5  import os
6
7
8  app = Flask(__name__)
9  CORS(app) # Permite conexión desde React
10
11
12 @app.route('/analizar', methods=['POST'])
13 def analizar():
14     data = request.get_json()
15     url = data.get('url', '')
16
17     if not url:
18         return jsonify({"error": "Falta la URL"}), 400
19
20     resultado = analizar_noticia(url)
21     return jsonify(resultado)
22
23
24 if __name__ == '__main__':
25     port = int(os.environ.get("PORT", 5000))
26     # Cambia a False en producción
27     app.run(host="0.0.0.0", port=port, debug=True)
28
29

```

En la ilustración 10 también podemos encontrar la configuración de los puertos y host para que la aplicación Render que es donde se encuentra desplegado la parte Back-End y el modelo funcionen.

10.3.3 Modelo de Clasificación

En el desarrollo del modelo de clasificación siendo esta la última fase del desarrollo de la aplicación web, tenemos un modelo creado en base a unos parámetros de evaluación basados únicamente en el texto referente a la noticia (este modelo no interpreta las imágenes o videos adjuntos a las noticias). Por medio de análisis de sentimientos y unas Apis de palabras claves se logra identificar la intención de los textos y dar un veredicto de la veracidad de las noticias. En las ilustraciones 11,12 y 13 podemos observar la composición del código y su forma de trabajar.

"Formando líderes para la construcción de un nuevo país en paz"



Ilustración 11 Composición principal del Modelo de Clasificación

```

1  # backend/model.py
2  import requests
3  from bs4 import BeautifulSoup
4  import re
5  import datetime
6  # from transformers import pipeline
7  import os
8
9  # classifier = pipeline("sentiment-analysis",
10 #                        model="distilbert-base-uncased-finetuned-sst-2-english")
11
12 # URLs de listas de palabras positivas y negativas en inglés y español
13 URL_POSITIVAS_EN = "https://ptrckprry.com/course/wordlist/positive.txt"
14 URL_NEGATIVAS_EN = "https://ptrckprry.com/course/wordlist/negative.txt"
15 URL_POSITIVAS_ES = "https://raw.githubusercontent.com/olea/lemarios/master/palabras-positivas.txt"
16 URL_NEGATIVAS_ES = "https://raw.githubusercontent.com/olea/lemarios/master/palabras-negativas.txt"
17
18
19 def cargar_lista_palabras(url, archivo_cache):
20     if os.path.exists(archivo_cache):
21         with open(archivo_cache, "r", encoding="utf-8") as f:
22             return [line.strip() for line in f if line.strip()]
23     else:
24         resp = requests.get(url)
25         palabras = [line.strip()
26                     for line in resp.text.splitlines() if line.strip()]
27         with open(archivo_cache, "w", encoding="utf-8") as f:
28             f.write("\n".join(palabras))
29         return palabras
30
31
32 palabras_positivas_en = cargar_lista_palabras(
33     URL_POSITIVAS_EN, "positivas_en.txt")
34 palabras_negativas_en = cargar_lista_palabras(
35     URL_NEGATIVAS_EN, "negativas_en.txt")
36 palabras_positivas_es = cargar_lista_palabras(
37     URL_POSITIVAS_ES, "positivas_es.txt")
38 palabras_negativas_es = cargar_lista_palabras(
39     URL_NEGATIVAS_ES, "negativas_es.txt")

```

En esta ilustración 11 podemos observar las primeras líneas de código del archivo del modelo de clasificación, en este podemos resaltar las líneas 9 y 10 las cuales son un comentario de la declaración de un modelo de análisis de sentimientos. Este modelo se utilizó para realizar pruebas de comparación en base al modelo desarrollado. El cual dio similitudes en las respuestas, por lo tanto, se corrobora el funcionamiento del modelo desarrollado.

Dejando esto de lado ya tenemos el inicio del modelo desarrollado, el cual para primero se consultan 4 APIs de palabras tanto en inglés como en español, para seguidamente encontrar una función para cargar los archivos .txt que serán los encargados de contener las palabras para comparar con las

"Formando líderes para la construcción de un nuevo país en paz"



noticias subidas. Ya en el final de la imagen tenemos el llamado a la función anterior por cada lista de palabras.

Ilustración 12 Composición secundaria del Modelo de Clasificación

```

1  alarm_words = [
2      "¡increíble!", "no lo vas a creer", "impactante", "última hora", "alerta", "escándalo"
3  ]
4
5
6  def detectar_idioma(texto):
7      # Detección simple: si hay muchas palabras con tildes, asume español
8      if re.search(r"[áéíóúñ]", texto.lower()):
9          return "es"
10     # Si hay palabras comunes en inglés
11     if re.search(r"\b(the|and|is|are|of|to|in|that|it|for|on|with)\b", texto.lower()):
12         return "en"
13     # Por defecto, español
14     return "es"
15
16
17 def analizar_sentimiento_simple(texto, idioma):
18     texto = texto.lower()
19     if idioma == "es":
20         negativos = sum(palabra in texto for palabra in palabras_negativas_es)
21         positivos = sum(palabra in texto for palabra in palabras_positivas_es)
22     else:
23         negativos = sum(palabra in texto for palabra in palabras_negativas_en)
24         positivos = sum(palabra in texto for palabra in palabras_positivas_en)
25     if negativos > positivos and negativos > 1:
26         return "NEGATIVE"
27     elif positivos > negativos and positivos > 1:
28         return "POSITIVE"
29     else:
30         return "NEUTRAL"
31
32
33 def analizar_noticia(url):
34     resultado = {
35         "fuente_confiable": False,
36         "buena_redaccion": False,
37         "fecha_actual": False,
38         "tiene_fuentes": False,
39         "sin_sensacionalismo": True,
40         "veredicto": "Análisis incompleto"
41     }

```

En la ilustración 12 podemos encontrar la continuación del archivo del modelo de clasificación, en esta empezamos con un vector de palabras alarmantes, que se utiliza para ayudar a filtrar algunos textos. Se nos presentan 3 funciones nuevas:

La primera función es utilizada para detectar el idioma del texto en base a unos parámetros específicos del idioma.



La segunda función es utilizada para el respectivo análisis del texto en base al idioma para poder utilizar el .txt necesario. Como ultima parte compara el total de palabras positivas y negativas para dar un veredicto.

La tercera función es la función principal la encargada de dar un resultado concreto de la noticia enviada por el usuario, y comienza declarando la estructura del objeto de retorna la respuesta final para mostrar en el Front-End.

Ilustración 13 Composición final del Modelo de Clasificación

```

1  try:
2      response = requests.get(url, timeout=10)
3      soup = BeautifulSoup(response.text, 'html.parser')
4      texto = soup.get_text(separator=' ', strip=True)
5
6      idioma = detectar_idioma(texto)
7
8      dominios_confiables = ["bbc.com", "cnn.com",
9                             "reuters.com", "elpais.com", "chequeado.com"]
10     resultado["fuente_confiable"] = any(
11         dominio in url for dominio in dominios_confiables)
12
13     errores = re.findall(
14         r'\b[kzxxq]{3,}|[^\a-zA-ZáéíóúÁÉÍÓÚñ\S0-9.,;:¿?¡()\'"]+', texto)
15     resultado["buena_redaccion"] = len(errores) < 5
16
17     fechas = re.findall(r'\d{1,2}/\d{1,2}/\d{2,4}', texto)
18     if fechas:
19         try:
20             fecha_detectada = datetime.datetime.strptime(
21                 fechas[0], '%d/%m/%Y')
22             resultado["fecha_actual"] = abs(
23                 (datetime.datetime.now() - fecha_detectada).days) < 90
24         except:
25             pass
26
27     resultado["tiene_fuentes"] = bool(
28         re.search(r"https?://\S+", texto)) or "según" in texto.lower()
29
30     resultado["sin_sensacionalismo"] = not any(
31         word in texto.lower() for word in alarm_words)
32
33     emocion = analizar_sentimiento_simple(texto[:500], idioma)
34     if emocion == "NEGATIVE":
35         resultado[

```



En la ilustración 13 podemos encontrar la estructura de la función principal analizar_noticia, en esta encontramos líneas para validar cada uno de los parámetros especificados en el objeto resultado, buscando dando respuesta a cada uno para lograr tener un veredicto acerca de la noticia.

10.4 Objetivo Especifico IV

Para dar cumplimiento al objetivo específico IV el cual consistía en validar la calidad para detectar Fake News de la aplicación web con IA basada en NLP, se utilizaron diferentes noticias de las paginas de noticias de Facebook. Paginas que habitualmente las personas consultan y consumen a diario.

Los resultados de estas pruebas los podemos observar en las ilustraciones 14,15,16 y 17.

Ilustración 14 Prueba 1

Detector Fake News

Información
No desconfíes más de la información de tus redes sociales.
Pega el link de la noticia y verifica su veracidad.

Link **Analizar**

Resultados

Parametros	resultado
Fuente confiable	SI
Buena redacción	SI
Fecha actual	NO
Tiene fuentes	NO
Sin sensacionalismo	SI

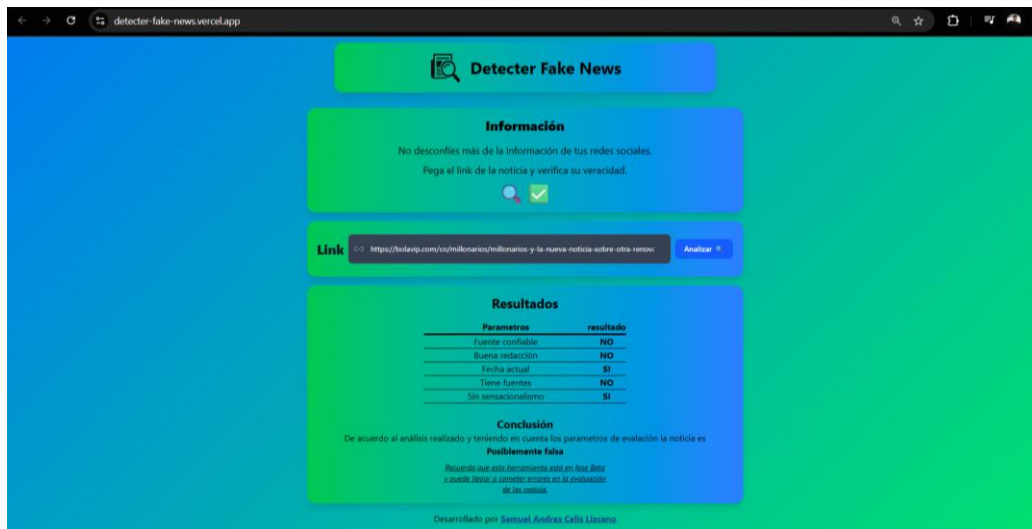
Conclusión
De acuerdo al análisis realizado y teniendo en cuenta los parametros de evaluación la noticia es
Posiblemente falsa
Recuerda que este herramienta solo es una guía y puede llevar a cometer errores en la evaluación de las noticias.

Desarrollado por Samuel Andres Cella Lázcano

Nota: Podemos observar que la primer prueba se realizó con un link de la pagina de enfoque Cúcuta jurídico, pagina dedicada a subir noticias de la ciudad de Cúcuta.

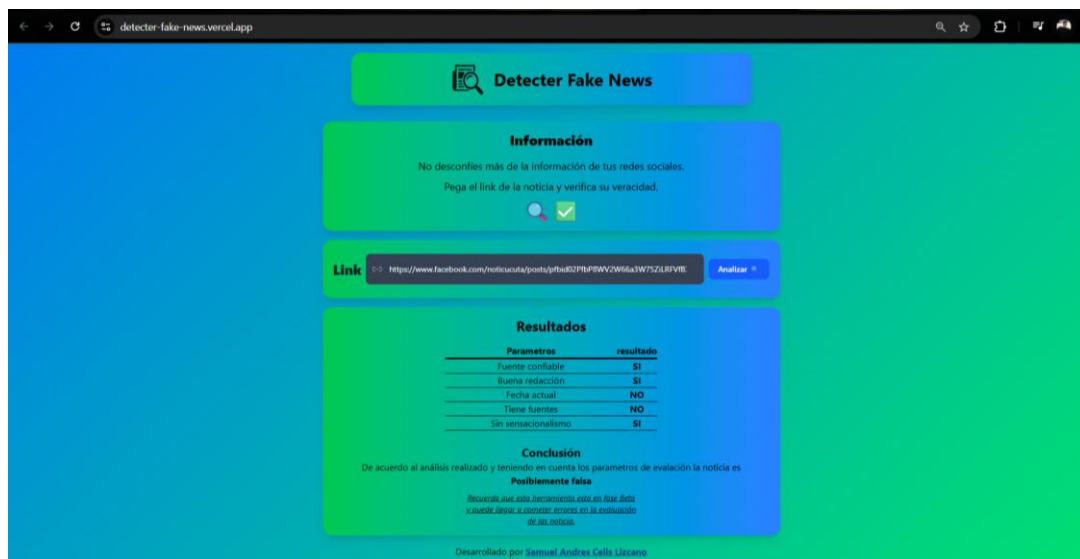


Ilustración 15 Prueba 2



Nota: En esta prueba podemos observar que el link probado es de un pagina llamada bolavip dedicada a subir noticias de futbol colombiano y todo lo relacionado a este.

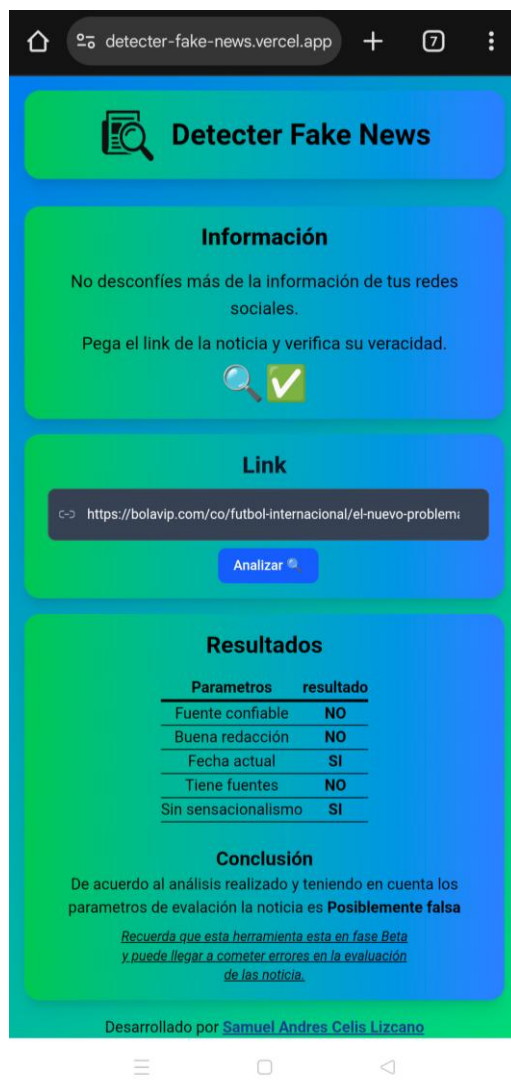
Ilustración 16 Prueba 3



Nota: En esta prueba podemos observar como el link utilizado es de una página llamada noticucuta, que se dedica a subir lo ultimo en noticias de la ciudad de Cúcuta.



Ilustración 17 Prueba 4



Nota: En esta última prueba se volvió a utilizar el link de una noticia de bolavip, pero desde un dispositivo móvil (celular).



11. Conclusiones y recomendaciones

- Durante la investigación de las herramientas y tecnologías necesarias para el desarrollo pudimos observar la gran variedad de opciones que existen actualmente en el mundo para llevar a cabo el desarrollo de un proyecto.
- En el momento de la selección de las herramientas y tecnologías necesarias para cada capa del desarrollo se logró comprender para que tipos de desarrollo son mejor cada una de las tecnologías existen, permitiendo así elegir la más apropiada con el objetivo general del proyecto.
- Al momento de construir aplicaciones web, el utilizar la arquitectura por capas el proyecto se simplifica de una forma gigantesca, permitiendo trabajar de manera simultanea en ambas partes del proyecto sin ver afectaciones en el desarrollo de alguna parte del proyecto.
- El desplegar los componentes de la aplicación en diferentes plataformas ayuda a mantener en funcionamiento los diferentes servicios de la aplicación.
- El realizar las pruebas de validación expuso los pocos dominios aceptados como validos en la lógica del modelo, permitiendo sugerencias como agregar las fuentes por medio de alguna Api que contenga los dominios de páginas oficiales.



12. Referencias

- 1) Wikipedia. (17 de 2 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
<https://es.wikipedia.org/wiki/CSS>
- 2) Agnieszka Jaff, K. U.-O.-M. (2020). Fake News and Aggregated Credibility: Conceptualizing a Co-Creative Medium for Evaluation of Sources Online. *International Journal of Ambient Computing and Intelligence*, 93-117.
- 3) Amirhosein Bodaghi, J. O. (2021). Thefakenewsgraphanalyzer:Anopen-sourcesoftwareforcharacterizing spreaders in large diffusion graphs. *Software Impacts*.
- 4) Ana Julia Dal Forno, G. P. (2025). Fake news detection algorithms – A systematic literature review. *Data & Knowledge Engineering*, 158.
- 5) Ariel José Torres Pineda, A. V. (2021). Veracidad y consumo de fake news entre estudiantes de una universidad pública de Cartagena de Indias, Colombia. *DOXA DIGITAL*, 153-174.
- 6) AWS. (12 de 5 de 2025). AWS. Obtenido de AWS: <https://aws.amazon.com/es/what-is/python/>
- 7) AWS. (19 de 5 de 2025). AWS. Obtenido de AWS: <https://aws.amazon.com/es/what-is/nlp/>
- 8) AWS. (19 de 5 de 2025). AWS. Obtenido de AWS: <https://aws.amazon.com/es/what-is/web-application/>
- 9) Castrillo, C. J. (2024). Paradigma Positivista. *Boletín Científico de las Ciencias Económico Administrativas del ICEA*, 29-32.
- 10) Chiung Chang Yu a, *. I. (2023). A Multi-tier Model and Filtering Approach to Detect Fake News Using . *INTERNATIONAL JOURNAL*, 643-651.



- 11) costanzo, m. (11 de 5 de 2025). *Platzi*. Obtenido de Platzi: <https://platzi.com/tutoriales/1248-pro-arquitectura/5439-patron-arquitectonico-de-capas-layers/#:~:text=La%20arquitectura%20en%20capas%20es,de%20lo%20que%20le%20cooresponde>).
- 12) Cuadros Bustamante, D. C. (2021). Análisis de fake news del periódico El Espectador y la Revista Semana sobre el covid-19, y el impacto que generan en los estudiantes de últimos semestres de Comunicación Social de la Universidad de Pamplona, sede Pamplona y la población sanalbertense. *Universidad de Pamplona - Facultad de Artes y Humanidades*.
- 13) Dave Bergmann, C. S. (23 de 11 de 2023). *IBM*. Obtenido de IBM: <https://www.ibm.com/es-es/topics/pytorch>
- 14) EAE. (12 de 5 de 2025). *EAE*. Obtenido de EAE Business School Barcelona: <https://www.eaebarcelona.com/es/blog/tensorflow>
- 15) EBAC. (1 de 6 de 2023). *EBAC Escuela Britanica de Artes Creativas y tecnología*. Obtenido de EBAC Escuela Britanica de Artes Creativas y tecnología: <https://ebac.mx/blog/que-es-html5#:~:text=HTML5%20es%20la%20%20C3%BAltima%20versi%C3%B3n,funcionalidad%20a%20los%20elementos%20HTML>.
- 16) google. (12 de 5 de 2025). *Google*. Obtenido de Visión general creada por IA: https://www.google.com/search?q=que+es+trello&sca_esv=b703f1cd0674f892&rlz=1C1UEAD_esCO1004CO1004&sxsrf=AHTn8zr0c7EgAH-5R8OPrBIiFqW0TXTdFQ%3A1747062971235&ei=uxAiaO6ODsGUwbkP_pe-yAU&ved=0ahUKEwui2LPenJ6NAXVBSjABHf6LD1kQ4dUDCBA&uact=5&oq=que+es+trello&gs_l



- 17) Google. (19 de 5 de 2025). *Google Cloud*. Obtenido de Google Cloud:
<https://cloud.google.com/learn/what-is-artificial-intelligence?hl=es-419>

- 18) Hugging face. (12 de 5 de 2025). *Hugging face*. Obtenido de Hugging face:
<https://huggingface.co/docs/transformers/index>

- 19) HugoQueiroz Abonizio, J. I. (2020). Language-Independent Fake News Detection: English, Portuguese, and Spanish Mutual Features. *Future internet*, 12-87.

- 20) IBM. (12 de 5 de 2025). *IBM*. Obtenido de IBM: <https://www.ibm.com/mx-es/topics/java-spring-boot>

- 21) IBM. (12 de 5 de 2025). *IBM*. Obtenido de IBM: <https://www.ibm.com/es-es/topics/mongodb>

- 22) IBM. (12 de 5 de 2025). *IBM*. Obtenido de IBM: <https://www.ibm.com/mx-es/topics/postgresql>

- 23) Issa Qiqieh, O. A.-Z. (2025). Anintelligent cyber threat detection: A swarm-optimized machine learning approach. *AlexandriaEngineeringJournal*, 553-563.

- 24) Javier. (24 de 8 de 2023). *FORMADORES (IT)*. Obtenido de FORMADORES (IT):
<https://formadoresit.es/que-es-postman-cuales-son-sus-principales-ventajas/>

- 25) KINSTA. (13 de 9 de 2023). *KINSTA*. Obtenido de KINSTA:
<https://kinsta.com/knowledgebase/what-is-express-js/>

- 26) KINSTA. (6 de 3 de 2025). *KINSTA*. Obtenido de KINSTA: <https://kinsta.com/es/base-de-conocimiento/que-es-node-js/>

- 27) López-Borrull, A. L.-B.-G. (2018). Fake news, ¿amenaza u oportunidad para los profesionales de la información y la documentación? *e-repository*.



- 28) mytaskpanel. (12 de 5 de 2025). *MYTASKPANEL*. Obtenido de MYTASKPANEL CONSULTING: <https://www.mytaskpanel.com/que-es-flask-beneficios-y-usos/#:~:text=Proporciona%20una%20manera%20sencilla%20de,preocupen%20por%20la%20infraestructura%20subyacente>.
- 29) OpenIA. (31 de 3 de 2025). *ChatGPT (version 31 de marzo)*. Obtenido de Obtenido de ChatGPT-4-turbo: <https://chatgpt.com/c/67cb48fc-9818-8004-ae56-b692e45a0647>
- 30) Pérez, C. R. (2019). No diga fake news, di desinformación: una revisión sobre el fenomeno de las noticias falsas y sus implicaciones. *COMUNICACIÓN*, 65 - 74.
- 31) Robledo, A. (24 de 9 de 2019). *OpenWebinars*. Obtenido de OpenWebinars: <https://openwebinars.net/blog/que-es-mysql/>
- 32) Rodríguez, C. A. (31 de 5 de 2020). *UNAM Global*. Obtenido de UNAM Global: https://unamglobal.unam.mx/global_revista/que-son-fake-news/
- 33) UNESCO. (6 de 3 de 2025). *UNESCO*. Obtenido de <https://www.unesco.org/es/seguimiento-regional-ods4>
- 34) Universidad de Alcalá. (12 de 5 de 2025). *Universidad de Alcalá*. Obtenido de Universidad de Alcalá: <https://www.master-data-scientist.com/pandas-herramienta-data-science/#:~:text=%C2%BFQu%C3%A9%20es%20pandas%3F,tablas%20num%C3%A9ricas%20y%20series%20temporales>.
- 35) W3School. (12 de 5 de 2025). *W3School*. Obtenido de W3School: https://www.w3schools.com/python/numpy/numpy_intro.asp#:~:text=NumPy%20is%20a%20Python%20library,you%20can%20use%20it%20freely.



- 36) Wikipedia. (16 de 9 de 2022). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
<https://es.wikipedia.org/wiki/SpaCy>
- 37) Wikipedia. (24 de 1 de 2024). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
<https://es.wikipedia.org/wiki/Bitbucket>
- 38) Wikipedia. (27 de 11 de 2024). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
<https://es.wikipedia.org/wiki/Jira>
- 39) Wikipedia. (19 de 11 de 2024). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
<https://es.wikipedia.org/wiki/React>
- 40) Wikipedia. (19 de 8 de 2024). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
[https://es.wikipedia.org/wiki/Angular_\(framework\)](https://es.wikipedia.org/wiki/Angular_(framework))
- 41) Wikipedia. (31 de 12 de 2024). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
<https://es.wikipedia.org/wiki/NLTK>
- 42) Wikipedia. (7 de 3 de 2025). *Wikipedia*. Obtenido de Wikipedia la enciclopedia libre:
<https://es.wikipedia.org/wiki/Modelo%20%93vista%20%93controlador>
- 43) Wikipedia. (10 de 5 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA La enciclopedia libre:
https://es.wikipedia.org/wiki/Visual_Studio_Code
- 44) Wikipedia. (22 de 4 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:
<https://es.wikipedia.org/wiki/PyCharm>
- 45) Wikipedia. (14 de 4 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA La enciclopedia libre:
<https://es.wikipedia.org/wiki/GitHub>



46) Wikipedia. (25 de 1 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA La enciclopedia libre:

<https://es.wikipedia.org/wiki/GitLab>

47) Wikipedia. (14 de 3 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:

<https://es.wikipedia.org/wiki/Figma>

48) Wikipedia. (7 de 1 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:

https://es.wikipedia.org/wiki/Tailwind_CSS

49) Wikipedia. (19 de 3 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA ka enciclopedia libre:

[https://es.wikipedia.org/wiki/Bootstrap_\(framework\)](https://es.wikipedia.org/wiki/Bootstrap_(framework))

50) Wikipedia. (10 de 5 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:

<https://es.wikipedia.org/wiki/JavaScript>

51) Wikipedia. (28 de 2 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:

<https://es.wikipedia.org/wiki/Vue.js>

52) Wikipedia. (21 de 4 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:

[https://es.wikipedia.org/wiki/Django_\(framework\)](https://es.wikipedia.org/wiki/Django_(framework))

53) Wikipedia. (12 de 5 de 2025). *WIKIPEDIA*. Obtenido de WIKIPEDIA la enciclopedia libre:

<https://en.wikipedia.org/wiki/Scikit-learn>

54) WIKIPEDIA. (16 de 3 de 2025). *Wikipedia La enciclopedia Libre*. Obtenido de Wikipedia La enciclopedia Libre: <https://es.wikipedia.org/wiki/Ontolog%C3%ADa>

55) Wikipedia. (10 de 4 de 224). *WIKIPEDIA*. Obtenido de WIKIPEDIA:

https://es.wikipedia.org/wiki/IntelliJ_IDEA