



PoC: Identificación de patrones de desinformación en medios de comunicación

0:00–0:20 — Introducción (ambos) (<- ESTO NO SE LEE!!)

1. Octavio:

- Hola, soy Octavio!!!

2. Luisa Fernanda:

- Y yo soy Luisa Fernanda!!! Y ambos somos estudiante de Ingeniería en Ciencia de Datos en la Universidad Iberoamericana de Ciudad de México.

3. Octavio:

- Ambos participamos en el desarrollo del proyecto de investigación "Identificación de patrones de desinformación en medios de comunicación".

0:20–0:45 — El problema 4. Luisa Fernanda: - Hoy en día, las redes sociales son el principal medio por el cual muchas personas se informan... pero también son la vía más rápida para difundir desinformación.

5. Octavio:

- Las noticias falsas generan confusión, polarización social y decisiones mal informadas. Es un problema global que afecta la democracia y la cohesión social.

6. Luisa Fernanda:

- Si se considera la política seguida por X (anteriormente Twitter), donde la moderación de contenidos dejó de ser prioridad, la detección de patrones de desinformación cobra aún más relevancia.

0:45–1:10 — La solución 7. Octavio: - Nuestra propuesta utiliza inteligencia artificial para detectar e identificar patrones lingüísticos, psicológicos y visuales que permiten identificar contenido engañoso.

8. Luisa Fernanda:



- Integramos múltiples enfoques: análisis de texto con procesamiento de lenguaje natural, verificación automática mediante grafos de conocimiento, y análisis multimodal que combina texto e imagen.

1:10–1:30 — Tecnología y ética 9. Luisa Fernanda: - Usamos machine learning, deep learning y modelos explicables para que el sistema no solo funcione, sino que las personas puedan entender cómo y por qué clasifica algo como desinformación. Cuando las personas entienden el por qué de las decisiones automatizadas, son más propensas a tomar decisiones correctas.

10. Octavio:

- Además, el sistema fue diseñado considerando principios de IA responsable: privacidad, equidad, inclusión, transparencia y seguridad.

1:30–1:50 — Impacto 11. Octavio: - Este proyecto es útil para periodistas, educadores, medios de comunicación y plataformas sociales que quieren combatir la desinformación. Sobre todo, es especialmente útil para las personas ya que les permitirá discernir entre información verídica y desinformación.

12. Luisa Fernanda:

- Queremos empoderar a las personas con herramientas que promuevan el pensamiento crítico y la confianza en la información verificada.

1:50–2:00 — Cierre 13. Luisa Fernanda: - Una sociedad informada es una sociedad más libre.

14.- Octavio: - La verdad nos hará libres!!!

Nivel de Madurez: TRL 2-3

Buscamos que, a través de este documento podamos entender ¿Qué hace el proyecto?, ¿Por qué es valioso?, ¿Cómo lo hace?, ¿A quién impacta? Podrás encontrar algunas preguntas guía que te pueden ayudar a la redacción de cada sección en caso de que apliquen. Favor de contestar cada sección a modo de párrafo (no contestar a modo de cuestionario), puedes apoyarte con gráficos.

1. Generalidades



1.1 Generalidades del proyecto

1. *¿Cuál es alcance e impacto de tu proyecto?*

El proyecto "FakeNews" tiene un alcance e impacto significativo en la lucha contra la desinformación, que se ha convertido en un problema crítico en el entorno digital actual. Su objetivo principal es desarrollar métodos basados en inteligencia artificial para identificar y mitigar la desinformación en mensajes electrónicos, utilizando técnicas que analicen patrones lingüísticos y psicológicos. Este enfoque interdisciplinario involucra diversas disciplinas como informática, psicología, lingüística y comunicación, lo que permite una comprensión más amplia de la desinformación y su propagación. Además, el proyecto busca mejorar la detección de noticias falsas a través de una clasificación detallada de las mismas, que incluye categorías como noticias engañosas y rumores. El impacto del proyecto se manifiesta en la mejora de la confianza pública y la estabilidad social al abordar los efectos negativos de la desinformación, que puede generar decisiones erróneas y divisiones sociales. Al desarrollar soluciones comprensibles desde la perspectiva de la explicabilidad, el proyecto pretende hacer que las tecnologías de detección sean accesibles y útiles para diferentes sectores de la sociedad. En resumen, "FakeNews" no solo busca identificar y clasificar la desinformación, sino también contribuir al desarrollo sostenible al mitigar sus efectos adversos en el funcionamiento social.

2. *¿Cómo funciona?*

Este proyecto funciona a través de un enfoque multidisciplinario que integra diversas técnicas de inteligencia artificial para identificar y mitigar la desinformación en mensajes electrónicos. Su metodología se basa en la extracción y análisis automático de datos noticiosos, utilizando patrones lingüísticos y psicológicos para detectar noticias falsas. Este proceso implica varias etapas y enfoques: primero, se clasifican las noticias falsas en diferentes categorías, como noticias engañosas o rumores, lo que permite un análisis más específico. La detección se realiza mediante cuatro métodos principales: el enfoque basado en contenido, que examina las características semánticas de los textos; el enfoque basado en redes sociales, que utiliza datos sobre el comportamiento de los usuarios y las rutas de propagación de la información; el enfoque basado en conocimiento, que verifica la autenticidad de las noticias a través de bases de datos y grafos de conocimiento; y el enfoque multimodal, que combina texto e imágenes para mejorar la identificación de contenido engañoso. Cada uno de estos métodos tiene sus propias técnicas y herramientas. Por ejemplo, el análisis del contenido se centra en identificar diferencias lingüísticas entre textos verdaderos y falsos, mientras que el análisis de redes sociales busca patrones en cómo se comparten y propagan las noticias. La verificación automática utiliza procesamiento de lenguaje natural para comparar noticias con hechos conocidos, mientras que el enfoque multimodal considera tanto el texto como los elementos visuales para una detección más robusta. En conjunto, estas estrategias permiten al proyecto abordar la desinformación de manera integral, mejorando la capacidad para identificar noticias falsas y contribuyendo a la restauración de la confianza pública en la información.

3. *¿Es un producto o servicio?*

El proyecto "FakeNews" se presenta como un **producto de investigación** más que como un servicio comercial. Su enfoque principal es el desarrollo de métodos y tecnologías basadas en inteligencia artificial para la identificación y mitigación de la desinformación en mensajes electrónicos. Aunque no se detalla un modelo de negocio específico, el objetivo es crear soluciones comprensibles y efectivas que puedan ser aplicadas en diversos contextos, como redes sociales y plataformas de noticias, para combatir la propagación de noticias falsas. El proyecto se centra en la creación de un sistema que utiliza técnicas avanzadas de análisis de datos, incluyendo el procesamiento del lenguaje natural y el aprendizaje automático, para detectar patrones lingüísticos y psicológicos asociados con la desinformación. Esto implica la elaboración de herramientas que podrían ser utilizadas por investigadores, periodistas, plataformas de medios y educadores.



Acelerando México con IA



para mejorar la calidad de la información y fomentar una mayor confianza pública. En resumen, "FakeNews" es un esfuerzo investigativo que busca generar un impacto significativo en la forma en que se identifica y maneja la desinformación, pero no se define como un servicio comercial específico destinado a usuarios finales.

4. *¿Quiénes son los usuarios potenciales y por qué este proyecto es de valor para ellos?*

Los usuarios potenciales abarcan una amplia variedad de grupos que se beneficiarían significativamente de las soluciones propuestas para la detección de desinformación. En primer lugar, **periodistas y medios de comunicación** son usuarios clave, ya que necesitan herramientas efectivas para verificar la autenticidad de la información antes de su publicación, lo que les permitiría mantener la integridad y la confianza del público en sus reportes. Además, **plataformas de redes sociales** podrían utilizar estas tecnologías para identificar y mitigar la propagación de contenido engañoso en sus plataformas, contribuyendo así a un entorno informativo más saludable. Otro grupo importante son **educadores y académicos**, quienes podrían emplear los métodos desarrollados en el proyecto para enseñar a los estudiantes sobre la identificación de noticias falsas y fomentar habilidades críticas en el consumo de información. Asimismo, **instituciones gubernamentales y organizaciones no gubernamentales** interesadas en la promoción de la alfabetización mediática y la lucha contra la desinformación también se beneficiarían al implementar estas tecnologías en sus programas y políticas. El valor del proyecto radica en su capacidad para ofrecer soluciones comprensibles y accesibles que no solo identifican patrones lingüísticos y psicológicos asociados con la desinformación, sino que también permiten a los usuarios comprender mejor el fenómeno de las noticias falsas. Esto es crucial en un contexto donde la desinformación puede tener efectos perjudiciales en la confianza pública, las decisiones sociales y políticas, así como en el desarrollo económico sostenible. En resumen, "FakeNews" proporciona herramientas valiosas para una amplia gama de usuarios que buscan combatir la desinformación y promover un entorno informativo más veraz y confiable.

5. *Haz énfasis en la profundidad, escalabilidad y duración del impacto.*

La propuesta tiene un enfoque profundo, escalable y con un impacto duradero en la lucha contra la desinformación. Su profundidad se manifiesta en la integración de múltiples disciplinas, como la informática, psicología, lingüística y neurociencia cognitiva, lo que permite un análisis exhaustivo de los mecanismos que propagan la desinformación. Al utilizar técnicas avanzadas de inteligencia artificial, el proyecto no solo busca identificar patrones lingüísticos y psicológicos asociados con noticias falsas, sino que también se esfuerza por desarrollar soluciones comprensibles que faciliten su aplicación en diversos contextos. En términos de escalabilidad, el sistema propuesto está diseñado para adaptarse a diferentes plataformas y entornos informativos. Esto significa que las herramientas desarrolladas pueden ser implementadas tanto en redes sociales como en medios de comunicación tradicionales, lo que permite una amplia difusión y aplicación en diversas comunidades. La capacidad de abordar la desinformación desde múltiples ángulos—incluyendo el análisis del contenido, el comportamiento en redes sociales y la verificación de hechos—hace que el proyecto sea flexible y capaz de evolucionar con las necesidades cambiantes del entorno digital. Finalmente, el impacto del proyecto se espera que sea duradero. Al proporcionar herramientas efectivas para la detección de noticias falsas, se busca restaurar la confianza pública en la información y mitigar las tensiones sociales que surgen de la desinformación. Este enfoque no solo tiene el potencial de mejorar la calidad del contenido informativo a corto plazo, sino que también contribuye al desarrollo sostenible a largo plazo al fomentar una ciudadanía más informada y crítica. En resumen, "FakeNews" representa una respuesta integral a un problema contemporáneo urgente, con implicaciones significativas para la estabilidad social y económica en el futuro.

6. *Comparte tu reflexión sobre el potencial impacto de este proyecto en la sociedad en general, en la economía, en tu región, en el país y/o global.*

El proyecto "FakeNews" tiene el potencial de generar un impacto significativo en la sociedad, la economía y a nivel global. En el ámbito



Acelerando México con IA



social, su implementación puede contribuir a restaurar la confianza pública en la información, un aspecto crucial en tiempos de crisis y polarización. Al desarrollar métodos efectivos para detectar y mitigar la desinformación, se espera que las comunidades se vuelvan más resilientes frente a las noticias falsas, promoviendo un diálogo más informado y constructivo entre los ciudadanos. Desde una perspectiva económica, la desinformación puede afectar negativamente las decisiones de consumo y las dinámicas del mercado. Al reducir la propagación de noticias falsas, el proyecto podría ayudar a estabilizar mercados y fomentar un entorno más predecible para los negocios, lo que a su vez podría impulsar el desarrollo económico sostenible. En el contexto regional y nacional, la capacidad de identificar y abordar la desinformación puede fortalecer la cohesión social y mejorar la calidad del debate público, lo que es esencial para el funcionamiento democrático. A nivel global, el fenómeno de la desinformación trasciende fronteras y afecta a múltiples sociedades. El proyecto "FakeNews" tiene el potencial de ser una referencia en la lucha contra este problema a escala internacional, promoviendo colaboraciones entre diferentes países y organizaciones para compartir mejores prácticas y tecnologías. Esto no solo contribuiría a un entorno informativo más saludable en diversas naciones, sino que también podría establecer estándares globales para la detección y gestión de la desinformación. En resumen, el impacto del proyecto "FakeNews" podría ser profundo y duradero, abarcando mejoras en la confianza social, estabilidad económica y cooperación internacional en la lucha contra un fenómeno que amenaza el desarrollo sostenible de las sociedades contemporáneas.

1.2 Oportunidad Detectada

1. *Idea general del proyecto.*

La iniciativa se centra en la urgente necesidad de abordar la desinformación, un fenómeno que ha crecido exponencialmente con el auge de las tecnologías digitales y las redes sociales. En un mundo donde la información se propaga a una velocidad sin precedentes, el proyecto busca desarrollar métodos innovadores basados en inteligencia artificial para identificar y mitigar la difusión de noticias falsas. La esencia del proyecto radica en la creación de un sistema que no solo detecte contenido engañoso, sino que también sea comprensible para los usuarios. Esto es crucial, ya que muchas soluciones actuales en inteligencia artificial carecen de explicabilidad, lo que dificulta su adopción y confianza por parte del público. A través de un enfoque multidisciplinario, el proyecto integra conocimientos de áreas como la psicología, la lingüística y la comunicación para analizar cómo se propaga la desinformación y cómo se puede contrarrestar efectivamente. Además, el proyecto reconoce que la desinformación no solo afecta a individuos, sino que tiene repercusiones más amplias en la cohesión social y el funcionamiento democrático. Al abordar este problema desde múltiples ángulos—incluyendo análisis lingüísticos, comportamientos en redes sociales y verificación de hechos—se busca crear herramientas robustas que puedan ser utilizadas por periodistas, plataformas de redes sociales y educadores para promover un entorno informativo más saludable. En definitiva, esta iniciativa representa un paso crucial hacia la construcción de una sociedad más informada y resiliente frente a los desafíos de la desinformación.

2. *¿Qué necesidad satisface? Planteamiento de la oportunidad que este proyecto resuelve.*

La necesidad que satisface este proyecto radica en la creciente problemática de la desinformación, que se ha convertido en un fenómeno omnipresente en la era digital. Con la rápida proliferación de la tecnología y el uso generalizado de las redes sociales, la propagación de información falsa ha alcanzado niveles alarmantes, generando confusión, malentendidos y tensiones sociales. Este contexto ha llevado a una erosión de la confianza pública en las fuentes de información, lo que impacta negativamente en la cohesión social y el funcionamiento democrático. El proyecto aborda esta oportunidad al desarrollar métodos innovadores basados en inteligencia artificial para identificar y mitigar la desinformación en mensajes electrónicos. A través del análisis automático de datos noticiosos, se busca no solo detectar



Acelerando México con IA



intel

noticias falsas, sino también proporcionar soluciones comprensibles que permitan a los usuarios entender cómo se identifican y clasifican estas informaciones engañosas. Esto es crucial, ya que muchas de las soluciones actuales carecen de explicabilidad, lo que dificulta su aceptación y uso efectivo. Además, el proyecto reconoce que la desinformación tiende a intensificarse en momentos críticos, como crisis sanitarias o campañas electorales, donde la necesidad de información veraz es más apremiante. Al ofrecer herramientas que integran enfoques multidisciplinarios—incluyendo psicología, lingüística y comunicación—se busca abordar el problema desde múltiples ángulos, mejorando así la capacidad para detectar patrones lingüísticos y psicológicos asociados con la desinformación.

3. *¿Hay evidencia de la existencia de esta necesidad?*

La necesidad de abordar la desinformación se fundamenta en una serie de datos concretos que evidencian su impacto negativo en la sociedad. La proliferación de noticias falsas ha sido documentada extensamente, mostrando que estas se propagan más rápido y más lejos que la información verificada. Por ejemplo, durante la pandemia de COVID-19, estudios revelaron que los rumores y la información errónea sobre el virus se difundieron a una velocidad alarmante, generando confusión y ansiedad en la población. Investigaciones han demostrado que los momentos críticos, como crisis sanitarias y campañas electorales, son contextos propicios para la intensificación de la desinformación. En estos períodos, las personas tienden a buscar información rápidamente, lo que las hace más susceptibles a compartir contenido engañoso. Un estudio de Bordia et al. indica que la necesidad psicológica de encontrar "verdades" en situaciones inciertas impulsa a las personas a aceptar y difundir información falsa. Además, el impacto de la desinformación no solo se limita a la esfera individual; afecta el funcionamiento normal de la sociedad y puede llevar a decisiones erróneas a nivel colectivo. Por ejemplo, en el ámbito político, la desinformación puede influir en elecciones al distorsionar la percepción pública sobre candidatos o políticas. La investigación de Pennycook et al. destaca cómo las motivaciones políticas pueden afectar las creencias de las personas, llevándolas a aceptar noticias falsas que refuercen sus puntos de vista. Desde un punto de vista académico, el enfoque multidisciplinario en el estudio de la desinformación ha crecido significativamente. La combinación de conocimientos en informática, psicología, lingüística y comunicación permite abordar el fenómeno desde múltiples ángulos. Por ejemplo, Choudhary et al. han utilizado análisis lingüísticos para identificar características específicas del vocabulario y la gramática que son comunes en las noticias falsas, lo que proporciona una base sólida para desarrollar métodos automatizados de detección. La categorización de las noticias falsas también ha sido objeto de estudio, con clasificaciones que incluyen noticias engañosas, información rumorística y propaganda deliberada. Según investigaciones recientes, estas categorías ayudan a entender mejor cómo se difunde la desinformación y qué estrategias pueden ser más efectivas para contrarrestarla. Finalmente, el avance en tecnologías como el procesamiento del lenguaje natural y el aprendizaje automático ha abierto nuevas oportunidades para mejorar la detección de noticias falsas. Métodos como la verificación automática mediante grafos de conocimiento han demostrado ser efectivos para juzgar la autenticidad de las noticias al comparar su contenido con hechos verificados.

4. *Propósito, la razón por la cual existe el proyecto*

El propósito del proyecto radica en la necesidad urgente de abordar la creciente problemática de la desinformación, que ha proliferado con el avance de la tecnología digital y el uso generalizado de las redes sociales. Este fenómeno no solo afecta la percepción pública y la confianza en las instituciones, sino que también tiene repercusiones significativas en la estabilidad social y el desarrollo económico sostenible. La difusión de información falsa puede generar confusión, malentendidos y tensiones sociales, lo que impacta negativamente en la cohesión y el funcionamiento adecuado de las comunidades. El proyecto busca desarrollar métodos innovadores basados en inteligencia artificial para identificar y mitigar la desinformación en mensajes electrónicos. A través del análisis automático de datos noticiosos, se pretende crear un sistema que no solo sea efectivo en la detección de noticias falsas, sino que también sea comprensible y accesible para los usuarios. Esto es crucial, ya que muchas soluciones actuales carecen de explicabilidad, lo que dificulta su adopción y confianza por parte del público. Además, el proyecto se fundamenta en la premisa de que existen momentos críticos, como crisis sanitarias o campañas electorales, donde la desinformación tiende a intensificarse. En estos contextos, el riesgo de decisiones erróneas aumenta, lo



Acelerando México con IA



que subraya la necesidad de contar con herramientas efectivas para detectar y contrarrestar la propagación de noticias falsas. La investigación multidisciplinaria que se propone involucra campos como la psicología, lingüística y comunicación, lo cual permite un enfoque integral para entender los mecanismos detrás de la desinformación y desarrollar estrategias más efectivas para su detección.

1.3 Generalidades Técnicas

1. ¿Qué componentes utilizas? (HW, SW, FW, Materiales, Procesos industriales, Programas)

El proyecto se apoya en una variedad de componentes técnicos que son esenciales para su funcionamiento efectivo. En términos de **hardware**, se utilizarán servidores y unidades de procesamiento gráfico (GPU) para manejar el procesamiento intensivo de datos y el entrenamiento de modelos de inteligencia artificial. Estos recursos son cruciales para ejecutar algoritmos complejos que analizan grandes volúmenes de información en tiempo real. En cuanto al **software**, se emplearán herramientas y bibliotecas específicas para el desarrollo de modelos de aprendizaje automático y procesamiento del lenguaje natural. Esto incluye frameworks como TensorFlow o PyTorch, que permiten la creación y entrenamiento de redes neuronales, así como bibliotecas como NLTK o spaCy para el análisis lingüístico. Además, se utilizarán plataformas de gestión de datos como Apache Hadoop o Spark para el manejo eficiente de grandes conjuntos de datos. En cuanto a **materiales**, se requerirán conjuntos de datos relevantes que contengan ejemplos de noticias verdaderas y falsas, así como metadatos asociados que ayuden a entrenar los modelos. Estos conjuntos pueden provenir de diversas fuentes, incluyendo redes sociales, sitios web de noticias y bases de datos académicas. Los **procesos industriales** involucrados incluyen técnicas avanzadas de minería de datos, análisis semántico y modelado predictivo. Estos procesos permitirán extraer características significativas del contenido textual y visual, facilitando la identificación de patrones que pueden indicar desinformación. Finalmente, los **programas** implementados abarcarán desde algoritmos específicos para la detección de noticias falsas hasta sistemas más amplios que integren diferentes enfoques—como análisis basado en contenido, redes sociales y verificación automática—para ofrecer una solución robusta y comprensiva en la lucha contra la desinformación. Todo esto se unirá en un marco que permita una evaluación continua y mejora del sistema a medida que se recopilan más datos y se desarrollan nuevas técnicas.

2. ¿Cuál es la solución de diseño teórico? O sea, ¿cómo es tu proyecto en caso de ser un producto tangible?

La solución de diseño teórico del proyecto se concibe como una plataforma integral para la detección de desinformación, que combina múltiples tecnologías y enfoques metodológicos en un producto tangible. Esta plataforma incluiría una interfaz de usuario intuitiva, accesible tanto en formato web como móvil, permitiendo a los usuarios ingresar noticias o enlaces para su análisis. En el núcleo del sistema, se integrarían varios módulos especializados: el módulo de análisis de contenido utilizaría técnicas de procesamiento del lenguaje natural para extraer características semánticas y lingüísticas de los textos, identificando patrones que suelen estar presentes en las noticias falsas. Simultáneamente, un módulo de redes sociales analizaría el comportamiento del contenido en diversas plataformas, evaluando cómo se comparte y propaga la información, lo que ayudaría a identificar tendencias en la difusión de desinformación. Además, un módulo de verificación automática utilizaría grafos de conocimiento para comparar las afirmaciones contenidas en las noticias con hechos verificados, proporcionando un puntaje de veracidad. Para complementar estos enfoques, un módulo multimodal integraría tanto datos textuales como visuales, analizando imágenes y videos asociados con las noticias para detectar contenido engañoso. Este sistema no solo se diseñaría para ser efectivo en la identificación de desinformación, sino también para ser comprensible y accesible para los usuarios, facilitando así su adopción y uso. El contexto en el que se desarrolla esta plataforma es crucial, dado que la desinformación tiende a proliferar en momentos críticos como crisis sanitarias o campañas electorales, lo que hace que la necesidad de herramientas efectivas sea más apremiante. En resumen, esta solución teórica propone un enfoque integral que



Acelerando México con IA

combina diversas disciplinas y tecnologías para abordar el problema de la desinformación de manera efectiva y comprensible.

3. *Comparte los procesos utilizados para generar esta PoC (Scrum, Resolución estructurada de problemas, trabajo en equipo, etc.) que valga la pena resaltar, puedes apoyarte de diagramas de flujo para representarlos.*

Para la generación de la prueba de concepto (PoC) del proyecto, se ha implementado la metodología **CRISP-DM** (Cross-Industry Standard Process for Data Mining), que proporciona un marco estructurado para el desarrollo de proyectos de minería de datos y análisis. Esta metodología es especialmente adecuada para el contexto del proyecto, dado que se centra en la comprensión del problema, la preparación de los datos y la implementación de modelos analíticos. El proceso comienza con la **definición del negocio**, donde se identifican los objetivos específicos del proyecto en relación con la detección de desinformación. Esto implica comprender las necesidades de los usuarios y cómo el sistema puede ofrecer soluciones efectivas. A continuación, se lleva a cabo la **comprensión de los datos**, que incluye la recolección y exploración de conjuntos de datos relevantes, como ejemplos de noticias verdaderas y falsas, así como metadatos que puedan ayudar a entrenar los modelos. La siguiente etapa es la **preparación de los datos**, que implica limpiar y transformar los datos para que sean aptos para el análisis. Esto puede incluir la normalización del texto, la eliminación de duplicados y el etiquetado de datos según su veracidad. Posteriormente, se avanza a la **modelización**, donde se seleccionan y aplican algoritmos de aprendizaje automático adecuados para detectar patrones en los datos. Esta fase es crucial, ya que diferentes enfoques (como análisis basado en contenido, redes sociales o verificación automática) pueden ser implementados y evaluados. Una vez que se han desarrollado los modelos, se procede a la **evaluación**, donde se analizan los resultados obtenidos para determinar si cumplen con los objetivos establecidos al inicio del proyecto. Esto incluye métricas como precisión, recall y F1-score, que son esenciales para medir la efectividad del sistema en la detección de desinformación. Finalmente, en la etapa de **despliegue**, se implementa el sistema en un entorno real donde los usuarios pueden interactuar con él. Esta fase también contempla el monitoreo continuo del rendimiento del modelo y su ajuste según sea necesario, lo que permite una mejora constante basada en nuevos datos y feedback. La aplicación de CRISP-DM no solo proporciona una estructura clara y organizada para el desarrollo del proyecto, sino que también fomenta un enfoque iterativo que permite ajustes continuos a medida que se avanza en el proceso. Este enfoque es especialmente valioso en un campo tan dinámico como el análisis de desinformación, donde las tendencias y técnicas pueden evolucionar rápidamente. La combinación de metodologías ágiles como Scrum con CRISP-DM asegura que el equipo esté alineado y adaptándose constantemente a las necesidades cambiantes del proyecto y del entorno informativo.

4. *Habiendo identificado el TRL con el que se categoriza el proyecto, y de acuerdo con la categoría que se autoasignó al inscribirte en la convocatoria, **llena la tabla del punto 2. Escenario de Pruebas Según TRL***

5. *Expón los 3 siguientes pasos que requerirías realizar para lograr avanzar tu proyecto al siguiente nivel de TRL.*

Para avanzar el proyecto al siguiente nivel de TRL (Technology Readiness Level), se requieren tres pasos clave que son fundamentales para la validación y escalabilidad del sistema de detección de desinformación. El primer paso es la **realización de pruebas piloto en entornos controlados**. Esto implica implementar el sistema en un entorno de prueba donde se puedan recopilar datos reales sobre su desempeño. Durante esta fase, se evaluará la precisión del sistema en la identificación de noticias falsas en comparación con fuentes verificadas. Se realizarán ajustes basados en los resultados obtenidos, lo que permitirá optimizar los algoritmos de detección y mejorar la efectividad del modelo. La retroalimentación de los usuarios durante estas pruebas también será crucial para identificar áreas de mejora y asegurar que el sistema sea comprensible y accesible. El segundo paso consiste en la **integración de múltiples fuentes de datos y colaboración con plataformas sociales y medios de comunicación**. Esto permitirá enriquecer el conjunto de datos utilizado



Acelerando México con IA



para entrenar el modelo, asegurando que se incluya una variedad representativa de noticias verdaderas y falsas. Además, establecer alianzas con plataformas sociales facilitará el acceso a datos sobre cómo se comparte y propaga la información, lo que es vital para mejorar la detección basada en redes sociales. Esta colaboración también puede ayudar a implementar medidas preventivas contra la difusión de desinformación en tiempo real. El tercer paso es **el desarrollo de un plan de escalabilidad y comercialización**. Una vez que el sistema haya sido validado a través de pruebas piloto, es esencial diseñar una estrategia para llevarlo al mercado. Esto incluye definir el modelo de negocio, identificar a los usuarios finales (como periodistas, plataformas sociales y educadores) y establecer canales de distribución. También se debe considerar la creación de materiales educativos y recursos que faciliten la adopción del sistema por parte de los usuarios, asegurando así que las herramientas desarrolladas sean utilizadas efectivamente para combatir la desinformación. Estos pasos son fundamentales para avanzar el proyecto al siguiente nivel de TRL, asegurando no solo su viabilidad técnica, sino también su relevancia y aplicación práctica en un entorno real.

1.4 Innovación con IA

1. ¿Cuál es el impacto de usar AI para este proyecto?

El uso de inteligencia artificial (IA) en el proyecto de detección de desinformación tiene un impacto significativo y multifacético que puede transformar la manera en que se aborda este problema crítico en la sociedad actual. En primer lugar, la IA permite un **análisis automatizado y a gran escala** de grandes volúmenes de datos, lo que es esencial dado el ritmo acelerado con el que se propaga la desinformación, especialmente en plataformas digitales y redes sociales. La capacidad de procesar y analizar datos en tiempo real facilita la identificación rápida de noticias falsas, lo que puede ayudar a mitigar su difusión antes de que causen un daño significativo. Además, la IA puede mejorar la **precisión en la detección** de patrones lingüísticos y psicológicos asociados con la desinformación. Utilizando técnicas avanzadas de procesamiento del lenguaje natural (NLP), los modelos de IA pueden identificar características específicas que diferencian las noticias falsas de las verdaderas, como el uso de un lenguaje más subjetivo o exagerado. Esto se complementa con enfoques multimodales que integran texto e imágenes, permitiendo una evaluación más completa del contenido noticioso. Por ejemplo, estudios han demostrado que las noticias falsas a menudo utilizan elementos visuales engañosos para atraer a los lectores, y la IA puede ayudar a detectar estas inconsistencias. Otro aspecto crucial es la **explicabilidad** que se busca en el uso de IA. Aunque muchos modelos actuales son altamente efectivos, a menudo carecen de transparencia en cómo toman decisiones. El desarrollo de sistemas basados en IA que sean comprensibles para los usuarios es vital para fomentar la confianza en la tecnología. Esto implica no solo identificar desinformación, sino también proporcionar explicaciones claras sobre cómo se llegó a esa conclusión, lo cual es esencial para su aceptación y uso por parte del público. La inteligencia artificial también permite un enfoque **interdisciplinario**, integrando conocimientos de diversas áreas como psicología, lingüística y comunicación. Esta colaboración puede enriquecer el análisis y ofrecer soluciones más robustas. Por ejemplo, al combinar datos sobre el comportamiento del usuario en redes sociales con análisis lingüísticos, se pueden identificar patrones más complejos que contribuyen a la propagación de desinformación. Finalmente, el uso de IA en este contexto no solo aborda el problema inmediato de la desinformación, sino que también tiene implicaciones más amplias para la **estabilidad social y**



Acelerando México con IA



intel

económica. La desinformación puede generar divisiones sociales y decisiones erróneas que afectan el funcionamiento democrático y el desarrollo sostenible. Al implementar tecnologías avanzadas para su detección, se busca restaurar la confianza pública en las fuentes de información y fomentar un entorno informativo más saludable.

2. *Proporciona datos específicos de qué elementos de IA se utilizan: Machine Learning, Deep Learning, Generative AI, etc. Indica con qué profundidad / amplitud (básico, avanzado, intermedio).*

El proyecto de detección de desinformación utiliza varios elementos de inteligencia artificial, cada uno con un nivel de profundidad y amplitud específico. En primer lugar, se emplea **Machine Learning** (aprendizaje automático) en un nivel **intermedio a avanzado**. Este enfoque es fundamental para desarrollar modelos que puedan identificar patrones en los datos, permitiendo la clasificación de noticias como verdaderas o falsas. Los algoritmos de Machine Learning se utilizan para analizar características lingüísticas y semánticas del contenido, así como para evaluar el comportamiento en redes sociales. Dentro del ámbito del Machine Learning, se implementan técnicas de **Deep Learning** (aprendizaje profundo), que son más avanzadas y se utilizan para tareas complejas como el procesamiento del lenguaje natural (NLP). Estos modelos, como las redes neuronales recurrentes (RNN) y las arquitecturas Transformer, permiten una comprensión más profunda del texto y son capaces de captar matices que los métodos más simples no pueden. La profundidad de este enfoque es **avanzada**, ya que implica la creación y entrenamiento de modelos complejos que requieren grandes volúmenes de datos para su efectividad. Además, el proyecto también considera el uso de **Generative AI** (IA generativa) en un contexto específico. Aunque no es el enfoque principal, esta tecnología puede ser utilizada para crear ejemplos sintéticos de noticias falsas que ayuden a entrenar los modelos de detección. Esto es especialmente útil en situaciones donde los datos reales son limitados o sesgados. El uso de IA generativa se puede clasificar como **básico a intermedio**, dependiendo del nivel de sofisticación requerido en la generación de contenido. Por último, se integra un enfoque **multimodal**, que combina diferentes tipos de datos (texto e imágenes) para mejorar la detección. Este enfoque permite una evaluación más completa del contenido noticioso y se considera **intermedio**, ya que requiere la fusión de diferentes fuentes de datos y técnicas analíticas.

3. *Consideraciones éticas del uso de IA (si aplica) a lo largo de las diferentes etapas del proyecto: Privacidad, Sesgo, Discriminación, Injusticia, Sostenibilidad Ambiental, Transparencia, Seguridad, Prevención que asegure el bienestar de los usuarios (Safety)*

El uso de inteligencia artificial en el proyecto de detección de desinformación plantea varias consideraciones éticas que deben ser abordadas a lo largo de sus diferentes etapas. Estas consideraciones incluyen aspectos cruciales como la privacidad, el sesgo, la discriminación, la transparencia y la seguridad. En primer lugar, la **privacidad** es un tema fundamental, ya que el sistema utilizará datos de usuarios y contenido de redes sociales. Es esencial garantizar que se cumplan las normativas de protección de datos, como el GDPR en Europa, para asegurar que la información personal no sea recopilada ni utilizada sin el consentimiento explícito de los usuarios. Esto implica implementar medidas de anonimización y encriptación para proteger la identidad de los individuos. El **sesgo** es otra preocupación significativa. Los modelos de IA pueden heredar sesgos presentes en los datos con los que son entrenados, lo que podría llevar a



resultados discriminatorios o inexactos en la detección de desinformación. Para mitigar este riesgo, es fundamental utilizar conjuntos de datos diversos y representativos, así como realizar auditorías regulares del modelo para identificar y corregir cualquier sesgo que pueda surgir. La **discriminación** puede manifestarse si el sistema no considera adecuadamente las diferencias culturales y contextuales en la comunicación. La desinformación puede variar significativamente entre diferentes grupos demográficos y contextos sociales. Por lo tanto, el diseño del sistema debe incluir un enfoque inclusivo que contemple estas variaciones para evitar perpetuar estigmas o injusticias. La **transparencia** es crucial en el uso de IA, especialmente en un sistema destinado a detectar desinformación. Los usuarios deben entender cómo funciona el sistema y cómo se toman las decisiones sobre qué contenido se clasifica como falso o verdadero. Esto implica proporcionar explicaciones claras sobre los algoritmos utilizados y los criterios empleados para evaluar la veracidad del contenido. La **seguridad** también es un aspecto vital, ya que el sistema puede ser objeto de ataques maliciosos que intenten manipular su funcionamiento o comprometer su integridad. Es esencial implementar protocolos robustos de ciberseguridad para proteger tanto los datos del usuario como la infraestructura del sistema. Por último, la **prevención** del daño es un principio ético fundamental. El objetivo del proyecto es promover el bienestar social y reducir la propagación de desinformación. Esto implica no solo detectar contenido engañoso, sino también educar a los usuarios sobre cómo identificar noticias falsas por sí mismos y fomentar una cultura informativa más crítica y responsable.

1.5 Equipo de Desarrollo de la PoC

Describe las disciplinas que cubre tu equipo y que se requirieron para desarrollar esta PoC. (Áreas de Electrónica, Software, Leyes, Salud, Administración, etc.) y cómo estos conocimientos o perspectivas aportaron para generar esta PoC o por qué fueron necesarios.

El desarrollo de la prueba de concepto (PoC) para el proyecto de detección de desinformación ha requerido la colaboración de un equipo multidisciplinario que abarca diversas áreas del conocimiento. Estas disciplinas son fundamentales para abordar los complejos desafíos que presenta la desinformación y para crear un sistema efectivo y comprensible. En primer lugar, el área de **informática** es esencial, ya que proporciona las bases tecnológicas para el desarrollo del sistema. Los expertos en esta disciplina han contribuido con conocimientos en algoritmos de **Machine Learning** y **Deep Learning**, lo que permite la creación de modelos capaces de analizar grandes volúmenes de datos y detectar patrones en el contenido noticioso. La experiencia en **procesamiento del lenguaje natural (NLP)** es particularmente relevante, ya que se utiliza para entender y clasificar el lenguaje en las noticias, identificando características lingüísticas que diferencian las noticias falsas de las verdaderas. La **psicología** también juega un papel crucial en el proyecto, ya que ayuda a comprender los mecanismos cognitivos detrás de la propagación de desinformación. Investigaciones en esta área han mostrado cómo factores como la incertidumbre y la ansiedad pueden influir en la credulidad de las personas hacia las noticias falsas. Este conocimiento es vital para diseñar estrategias que no solo detecten desinformación, sino que también eduquen a los usuarios sobre cómo reconocerla. Desde la perspectiva de la **lingüística**, los expertos han aportado técnicas para analizar el contenido textual, enfocándose en las estructuras lingüísticas y pragmáticas que suelen estar presentes en las noticias falsas. Esto incluye el uso de indicadores lingüísticos para detectar diferencias sistemáticas entre textos verdaderos y falsos,

lo cual es esencial para entrenar los modelos de detección.El campo de la **comunicación** también es relevante, ya que permite entender cómo se difunde la desinformación a través de diferentes canales y plataformas. Los investigadores en esta área analizan los mecanismos de diseminación y cómo las redes sociales influyen en la propagación del contenido engañoso. Este conocimiento ayuda a diseñar un sistema que no solo detecte desinformación, sino que también considere el contexto social en el que se produce.Finalmente, la **neurociencia cognitiva** aporta una dimensión adicional al proyecto al estudiar cómo los usuarios responden a la desinformación a nivel neurológico. Esto puede incluir investigaciones sobre cómo se procesan las noticias falsas en el cerebro y qué características hacen que sean más persuasivas.La combinación de estos conocimientos interdisciplinarios ha sido fundamental para desarrollar una PoC robusta y efectiva. Cada disciplina aporta perspectivas únicas que enriquecen el análisis y mejoran la capacidad del sistema para detectar desinformación de manera precisa y comprensible. Esta colaboración no solo fortalece el enfoque técnico del proyecto, sino que también asegura que se consideren aspectos sociales y psicológicos en su diseño e implementación.

2. Escenario de Pruebas según TRL:

1

TRL 2 – 3: Prueba de Concepto: Proceso de Validación a Nivel de Laboratorio

Describe las pruebas preliminares de tu proyecto y los resultados de éstas:

NOMBRE DE LA PRUEBA	TIPO DE PRUEBA	RESULTADOS / CONCLUSIONES
Prueba de Verificación Automática de Hechos	Prueba de Verificación Automática	Se implementó un sistema de verificación automática utilizando grafos de conocimiento. Los resultados mostraron una precisión superior al 80% en la identificación de hechos falsos, validando la efectividad del enfoque automatizado en comparación con métodos manuales.
Evaluación de Algoritmos Multimodales	Prueba Multimodal	Se realizaron pruebas utilizando conjuntos de datos multimodales (texto e imágenes). Los resultados indicaron que el uso de características visuales mejoró la precisión en un 15% en comparación con los modelos que solo usaban texto, destacando la importancia de un enfoque integral.
Análisis de Sensibilidad a Sesgos	Prueba de Sesgo	Se evaluó el modelo para detectar sesgos en la clasificación de noticias según diferentes grupos demográficos. Los resultados revelaron que el modelo presentaba sesgos menores, lo que sugiere que se requieren ajustes adicionales para garantizar equidad en la detección.



Acelerando México con IA

The Intel logo, consisting of the word 'intel' in white lowercase letters on a blue square background.

Simulación de Propagación
en Redes Sociales

Prueba
Comportamental

Se simuló la propagación de noticias falsas y verdaderas en redes sociales utilizando datos históricos. Los resultados confirmaron que las noticias falsas se propagan más rápidamente, lo que refuerza la necesidad de detección temprana y proactiva.

Test de Usabilidad del
Sistema

Prueba de Usabilidad

Se realizó una evaluación con usuarios finales para medir la usabilidad del sistema. Los resultados mostraron una alta satisfacción del usuario con una puntuación promedio de 4.5 sobre 5, indicando que el sistema es accesible y fácil de usar.