# OWNML MACHINE LEARNING CANVAS Designed for: Designed by: Date: Iteration: .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TAREA DE PREDICCIÓN **¿Cuál es el tipo de aprendizaje?**  Aprendizaje supervisado  **¿Si es aprendizaje supervisado, indicar qué se predice?**  Se predice si una noticia es "Falsa" (noticia falsa) o "Real" (noticia auténtica)  **¿Cuáles son los posibles resultados de la tarea de aprendizaje?**  "Falsa" (noticia falsa) o "Real" (noticia auténtica)  **¿Cuándo se observan los resultados de esta tarea?**  Los resultados se obtienen en tiempo real cuando un usuario analiza un artículo, generalmente en segundos. | DESICIONES **¿Cómo se convierten las predicciones en recomendaciones o decisiones accionables para el usuario final?**  Marcar artículos como potencialmente falsos para su revisión por verificadores de hechos  Proporcionar puntuaciones de confianza sobre la autenticidad de las noticias  Informar a los usuarios a través de aplicaciones web, extensiones de navegador o API | PROPUESTA DE VALOR**¿Quién es el beneficiario final?**Periodistas, responsables políticos y el público en general.**¿De qué empresa es?**Organizaciones de verificación de hechos, medios de comunicación y plataformas digitales que buscan combatir la desinformación.**¿Qué problemas específicos se abordan?**Difusión de noticias falsas que afectan la percepción pública y las decisiones políticas.Manipulación informativa que puede influir en procesos electorales y en la gobernabilidad.Falta de herramientas accesibles para verificar información en tiempo real.**¿Qué riesgo puede tener para ese beneficiario el uso de este modelo?**Posibles falsos positivos que etiqueten noticias verídicas como falsas, afectando la credibilidad de medios legítimos.Sesgo en la detección que pueda favorecer o perjudicar ciertos grupos políticos o ideológicos.Dependencia excesiva en la herramienta sin verificación humana adicional, lo que podría llevar a errores en la interpretación de los resultados. | RECOLECCIÓN DE DATOS – NO SE DEBE DILIGENCIAR | FUENTES DE DATOS **¿Qué fuentes de datos se utilizan?**  Bases de datos internas de verificadores de hechos con etiquetas previas de noticias.  APIs de periódicos digitales y sitios de noticias.  Archivos históricos de medios de comunicación.  Conjuntos de datos públicos sobre noticias falsas y reales.  **¿De dónde se toman los datos?**  Se extraen de sitios web de noticias, APIs de fuentes verificadas y bases de datos de fact-checking.  Se recopilan mediante web scraping y colaboración con plataformas de medios.  **¿Se pueden utilizar para realizar el objetivo del análisis?**  Sí, los datos han sido seleccionados y procesados para garantizar su validez y aplicabilidad en la detección de noticias falsas.  Se aplican técnicas de limpieza y normalización para mejorar la calidad del análisis. |
| SIMULACIÓN DE IMPACTO **¿Cuáles son los valores de costo/beneficio de las decisiones (in)correctas?**  Decisiones correctas: Reducción de la propagación de noticias falsas, mayor confianza en los medios de comunicación, fortalecimiento de la democracia.  Decisiones incorrectas: Etiquetado erróneo de noticias verídicas como falsas, disminución de la credibilidad de los medios de verificación, posible censura involuntaria de información legítima.  **¿Cuáles son los criterios de éxito del modelo para su posterior despliegue?**  Alcanzar un umbral de precisión mínimo del 90% en pruebas controladas.  Evaluación positiva en múltiples idiomas y estilos periodísticos.  Capacidad de adaptación a nuevas estrategias de desinformación mediante aprendizaje continuo.  **¿Existen restricciones de equidad?**  Evitar sesgos ideológicos que favorezcan o perjudiquen a ciertos grupos políticos.  Garantizar una evaluación equitativa de noticias provenientes de diversas fuentes.  Implementar auditorías periódicas para detectar posibles discriminaciones en la detección de noticias falsas. | APRENDIZAJE (USO DEL MODELO) **¿El uso del modelo es por lotes o en tiempo real?**El modelo se usa en tiempo real para analizar noticias conforme son procesadas por los usuarios.**¿Con qué frecuencia se usa?**Se utiliza cada vez que un usuario envía un artículo para verificación, permitiendo una detección inmediata de noticias falsas. |  | CONSTRUCCIÓN DE MODELOS**¿Cuántos modelos se necesitan?**Se requiere un modelo principal con posibles variaciones o submodelos adaptados a distintas fuentes de noticias. Pero para este caso se solicitan realizar 3 modelos distintos donde se haga una comparativa cual es el más óptimo para identificar los requerimientos.**¿Cuándo deben actualizarse?**Se deben actualizar mensualmente con nuevos datos etiquetados para mejorar su precisión y adaptación a cambios en la generación de noticias falsas.**¿De cuánto tiempo se dispone para generar el modelo (incluido el proceso de ingeniería de características y el análisis o evaluación del mismo)?**El proceso de construcción del modelo, incluyendo la ingeniería de características y la evaluación, debe completarse en un período de 2 a 4 semanas antes de cada actualización mensual. | INGENIERÍA DE CARACTERÍSTICAS**¿Qué variables/características se utilizan en el modelo?** Frecuencia de palabras clave asociadas a noticias falsas.Análisis de estructura gramatical y sintáctica del texto.Metadatos de las noticias, como la fuente y fecha de publicación.**¿Qué agregaciones o transformaciones se aplican a las fuentes de datos originales (incluir las más importantes)?** Vectorización mediante TF-IDF para ponderar la importancia de términos.Análisis de sentimiento para evaluar el tono del contenido.Normalización de texto eliminando caracteres especiales y aplicando lematización. |
|  | MONITOREO -NO SE DEBE DILIGENCIAR |  |  |  |

|  |  |  |
| --- | --- | --- |
|  | Version 1.2. Created by Louis Dorard, Ph.D. Licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/). Please keep this mention and the link to [ownml.co](https://www.ownml.co/) when sharing. | [**OWNML.CO**](https://www.ownml.co/) |

# 