










<p><b>TAREA DE APRENDIZAJE</b></p>  <p>La tarea de aprendizaje es supervisada, en la cual se pretende predecir la veracidad de una noticia. Es una tarea de clasificación binaria que pretende establecer si la noticia es “verdadera” o “falsa”. Se espera que el modelo idealmente genere la predicción en tiempo real o en intervalos cortos de tiempo de no más unas cuantas horas, para que permita una detección rápida y así evitar la propagación de desinformación.</p>	<p> <b>DECISIONES</b></p> <p>Si el modelo identifica una noticia con alta probabilidad de ser falsa, se puede etiquetar en la plataforma a la que pertenece o enviar una alerta al equipo de verificación.</p> <p>En plataformas de redes sociales, la noticia podría ser puesta en revisión o su difusión podría ser limitada hasta una verificación manual.</p> <p>Proveer una nota informativa o una advertencia sobre la posible falsedad del contenido, permitiendo al usuario tomar una decisión informada.</p>	<p> <b>PROPUESTA DE VALOR</b></p> <p>Los beneficiarios finales serían los ciudadanos del común ya que podría bajar el consumo de desinformación y proteger la integridad de la opinión pública en pro de una mejor toma de decisiones en su participación política y pública. Así mismo las plataformas de información y redes sociales se beneficiarían de este modelo al mejorar la calidad del contenido difundido y evitar la manipulación de la información, cumpliendo su compromiso social y ofreciendo una propuesta de valor adicional a sus usuarios. Por último, los organismos gubernamentales son también beneficiarios ya que puede ser un instrumento de apoyo en la toma de decisiones y la estabilidad en procesos electorales y diplomáticos.</p>	<p> <b>RECOLECCIÓN DE DATOS – NO SE DEBE DILIGENCIAR</b></p> <p>¿Cómo se obtiene el conjunto inicial de entidades y resultados (por ejemplo, extractos de bases de datos, extracciones de API, etiquetado manual)? ¿Qué estrategias se aplican para actualizar los datos continuamente, controlando los costos y manteniendo la vigencia?</p>	<p> <b>FUENTES DE DATOS</b></p> <p>Los datos vienen de periódicos en línea como ‘Público’, ‘La Marea’ y ‘El Común’ facilitados por la Universidad de los Andes como recurso del curso Inteligencia de Negocios del departamento de Sistemas y Computación. Los datos son otorgados bajo licencia Attribution 4.0 International</p>
<p> <b>SIMULACIÓN DE IMPACTO</b></p> <p><b>Beneficio:</b> Reducción de la difusión de desinformación, fortalecimiento de la confianza en las instituciones y mejora en la calidad informativa.</p> <p><b>Costo:</b> Consecuencias negativas derivadas de errores en la clasificación (falsos positivos y falsos negativos), que pueden afectar la reputación de la plataforma o la percepción pública.</p> <p><b>Criterios de éxito:</b></p> <p>Alto recall y F1-score en la detección</p> <p>Tiempo de respuesta cortos que permitan una intervención oportuna.</p> <p>- Aceptación y confianza de los usuarios y moderadores.</p>	<p> <b>APRENDIZAJE (USO DEL MODELO)</b></p> <p>El uso del modelo se pretende que sea en tiempo real, se ejecutaría de forma continua o, si es el caso, en intervalos cortos de tiempo para asegurar la detección oportuna y minimizar el riesgo de propagación de noticias falsas.</p>	<p>Los problemas que se abordan son:</p> <ul style="list-style-type: none"> <li>Mitigación de la polarización y manipulación de la opinión pública.</li> <li>Prevención de riesgos en procesos electorales y tensiones internacionales.</li> </ul> <p>Dependiendo de la efectividad del modelo y su uso responsable los riesgos a los que podría un beneficiario puede verse expuesto por su uso son. <b>Riesgos para el beneficiario:</b></p> <ul style="list-style-type: none"> <li>Que se etiqueten incorrectamente noticias veraces, lo que podría afectar la libertad de expresión y la credibilidad de la plataforma.</li> <li>Que no se detecten noticias falsas, permitiendo su propagación y sus consecuencias negativas.</li> </ul> <p>Riesgo de que el modelo adquiera o perpetúe sesgos ideológicos, afectando la imparcialidad en la detección.</p>	<p> <b>CONSTRUCCIÓN DE MODELOS</b></p> <p>En principio se pretende construir 3 modelos para evaluar su eficacia en la detección de noticias falsas y escoger uno principal, sin embargo, luego se podría evaluar si es necesario la combinación de modelos para obtener mejores resultados.</p> <p>La actualización del modelo escogido tendría que ser periódica con el fin de incorporar nuevos datos para mantener la precisión del modelo ante el cambio de temas o tendencias informativas.</p> <p>Se cuenta con 3 semanas para el levantamiento y evaluación de los modelos.</p>	<p> <b>INGENIERÍA DE CARACTERÍSTICAS</b></p> <p>Se pretende trabajar con el título y descripción de las noticias proporcionadas.</p> <p>Estos textos serán pasados por un proceso de limpieza que consiste en extraer las palabras y símbolos poco significantes para el modelo, como lo son la puntuación, los artículos, los conectores, entre otros. Así luego se pasa por un proceso de tokenización, donde cada palabra va a representar un token en nuestro modelo.</p> <p>A la lista de estos tokens se les realiza un proceso de vectorización con TF-IDF con el fin de tener datos numéricos aptos para el entrenamiento de modelos.</p>

<p><b>Restricciones de equidad:</b></p> <p>El modelo debe garantizar neutralidad, evitando sesgos políticos o ideológicos que puedan favorecer o perjudicar a determinados grupos.</p>				
	<p><b>MONITOREO NO SE DEBE DILIGENCIAR</b></p> <p>¿Qué métricas y KPI se utilizan para hacer un seguimiento del impacto de la solución de ML una vez desplegada, tanto para los usuarios finales como para la empresa?</p> <p>¿Con qué frecuencia deben revisarse?</p>			

## Adaptación de OWNML MACHINE LEARNING CANVAS



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://ownml.co) when sharing.

**[OWNML.CO](https://ownml.co)**