











<p><b>TAREA DE APRENDIZAJE</b></p> <p></p> <p>¿Cuál es el tipo de aprendizaje?</p> <p>El tipo de aprendizaje es supervisado con clasificación multiclase.</p> <p>¿Si es aprendizaje supervisado, indicar qué se predice?</p> <p>Se busca predecir el ODS al que pertenece a una opinión ciudadana, donde si ODS 1 es el fin de la pobreza, el ODS 3 la salud y bienestar, y el ODS 4 la educación de calidad.</p> <p>¿Cuáles son los posibles resultados de la tarea de aprendizaje?</p> <p>La etiqueta correspondiente al ODS asociado al texto.</p> <p>¿Cuándo se observan los resultados de esta</p>	<p> <b>DECISIONES</b></p> <p>¿Cómo se convierten los resultados del modelo en recomendaciones o decisiones procesables para el usuario final?</p> <p>Los resultados se convierten en clasificaciones procesables para que analistas y responsables de política pública puedan identificar las principales problemáticas que mencionan los ciudadanos. Orientar recursos y programas hacia los ODS más mencionados. Priorizar acciones de intervención con evidencia de opinión pública.</p>	<p> <b>PROPUESTA DE VALOR</b></p> <p>¿Quién es el beneficiario final? ¿De qué empresa es?</p> <p>El beneficiario final es el fondo de población de las Naciones Unidas (UNFPA) y entidades gubernamentales que hacen seguimiento a la Agenda 2030. La dificultad de procesar grandes volúmenes de opiniones ciudadanas y clasificarlas en los ODS relevantes.</p> <p>¿Qué problemas específicos se abordan?</p> <p>La dificultad de procesar grandes volúmenes de opiniones ciudadanas y clasificarlas en los ODS relevantes.</p> <p>¿Qué riesgo puede tener para ese beneficiario el uso de este modelo?</p> <p>Los riesgos que puede tener para este beneficiario el uso de este modelo son: Clasificaciones erróneas que lleven a priorizar un ODS</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>¿Qué fuentes de datos se utilizan? (Mencione tablas de bases de datos internas y externas o métodos API).</p> <p>Se emplean principalmente los archivos Excel suministrados por los docentes, que contienen opiniones ciudadanas en texto libre junto con su respectiva etiqueta (ODS 1, ODS 3 u ODS 4). Estas tablas constituyen la base de datos interna inicial para el entrenamiento y prueba del modelo. En fases posteriores, la solución podría integrarse con APIs de participación ciudadana (por ejemplo, plataformas gubernamentales o del UNFPA), así como con bases de datos abiertas de encuestas y consultas públicas que recolectan opiniones en torno a los Objetivos de Desarrollo Sostenible.</p> <p>¿De dónde se toman los datos?</p> <p>Los datos iniciales provienen directamente del curso y del enunciado del proyecto, con el fin de garantizar un conjunto de entrenamiento representativo. Futuras integraciones podrían conectarse a repositorios institucionales del UNFPA y a datos no estructurados de redes sociales o foros ciudadanos.</p> <p>¿Se pueden utilizar para realizar el objetivo del análisis?</p> <p>Sí, los datos pueden utilizarse para el objetivo del análisis, dado que cuentan con la variable independiente (texto en lenguaje natural) y la variable objetivo (ODS asociado). Esto permite abordar el problema como una tarea de clasificación supervisada multiclase. Además, su formato estructurado (tablas Excel) facilita el preprocesamiento y la transformación a representaciones vectoriales para el modelo de machine learning.</p>
--	--	--	--	---

<p>tarea? Por ejemplo, si es un modelo predictivo indicar si el resultado se obtiene unas horas, días, semanas, o meses antes.</p> <p>El modelo entrega la clasificación inmediatamente después de recibir un nuevo texto, por lo que la predicción es en segundos.</p>		<p>equivocado. Sesgos en el modelo que subrepresenten problemáticas de ciertos grupos. Riesgo de confianza institucional si las predicciones no son transparentes.</p>		
<p> SIMULACIÓN DE IMPACTO</p> <p>¿Cuáles son los valores de costo/beneficio de las decisiones (in)correctas?</p> <p>FP = 1x. Opción mal asignada, re-trabajo.</p> <p>FN en la clase prioritaria = 2x. Se pierde.</p>	<p> APRENDIZAJE (USO DEL MODELO)</p> <p>¿El uso del modelo es por lotes o en tiempo real?</p> <p>Por lotes diarios. En tiempo real solo para picos, es decir, colas de mensajes.</p> <p>¿Con qué frecuencia se usa?</p>		<p> CONSTRUCCIÓN DE MODELOS</p> <p>¿Cuántos modelos se necesitan?</p> <p>Tres modelos comparables. Naive-Bayes, KNN y regresión logística.</p> <p>¿Cuándo deben actualizarse?</p>	<p> INGENIERÍA DE CARACTERÍSTICAS</p> <p>¿Qué variables/características se utilizan en el modelo?</p> <p>TF-IDF de palabras (1-2-gram) con min_df=2-5.</p> <p>Metadatos simples como longitud del texto, presencia de números/numerales si aportan.</p> <p>¿Qué agregaciones o transformaciones se aplican a las fuentes de datos originales – incluir las más importantes--?</p> <p>Normalización en minúsculas y tildes, limpieza de URLs, menciones y emoticones.</p>

<p>Acuerdo = 0x.</p> <p>¿Cuáles son los criterios de éxito del modelo para su posterior despliegue?</p> <p>F1-macro <math>\geq 0.75</math> y Balanced Accuracy <math>\geq 0.75</math> con partición de prueba.</p> <p>¿Existen restricciones de equidad?</p> <p>Monitorear métricas por fuente, dialecto y longitud. Anonimizar identificadores, auditorías trimestrales de sesgo.</p>	<p>Se usa diariamente. Se infiere bajo demanda cuando llegan opiniones nuevas.</p>		<p>Mensualmente o si F1-macro <math>&gt; 5</math> p.p.</p> <p>¿De cuánto tiempo se dispone para generar el modelo (incluido el proceso de ingeniería de características y el análisis o evaluación del mismo)?</p> <p>Entrenamiento + evaluación <math>\leq 30-60</math> minutos con tabla comparativa. Se promueve a producción únicamente si se superan los umbrales.</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? ¿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-nc-sa/4.0/).

Please keep this mention and the link to [ownml.co](https://ownml.co) when sharing.

**OWNML.CO**