






<p>TAREA DE APRENDIZAJE</p> <p></p> <p>¿Cuál es el tipo de aprendizaje? ¿Si es aprendizaje supervisado, indicar qué se predice? ¿Cuáles son los posibles resultados de la tarea de aprendizaje? ¿Cuándo se observan los resultados de esta tarea?</p> <p>El aprendizaje es de tipo supervisado. Se predice que objetivo de desarrollo sostenible está asociado a la opinión dada. La tarea clasifica la opinión en tres objetivos de desarrollo sostenible (ODS 1, 3 Y 4). Los resultados se obtienen de manera inmediata al momento de inferencia.</p>	<p> DECISIONES</p> <p>¿Cómo se convierten los resultados del modelo en recomendaciones o decisiones procesables para el usuario final?</p> <p>Los resultados del modelo se convierten en recomendaciones procesables al clasificar cada opinión de los ciudadanos dentro de los ODS 1, 3 y 4. Este resultado es valioso para el usuario final (agencias vinculadas a los ODS), ya que le entrega información para realizar acciones concretas a partir de la predicción (clasificación).</p>	<p> PROPUESTA DE VALOR</p> <p>¿Quién es el beneficiario final? ¿De qué empresa es? ¿Qué problemas específicos se abordan? ¿Qué riesgo puede tener para ese beneficiario el uso de este modelo?</p> <p>El beneficiario final son los ciudadanos, ya que las decisiones que tomen las entidades, respecto a la pobreza, salud y educación les afectan a ellos.</p> <p>Los problemas que se abordan son el alto volumen de datos no estructurados y la dificultad que puede generar vincular las opiniones con los ODS.</p> <p>El uso de este modelo puede ser un riesgo para el beneficiario final cuando haya errores de clasificación (falsos positivos o falsos negativos). Esto puede llevar a que las entidades tomen decisiones incorrectas basadas en el modelo.</p>	<p> RECOLECCIÓN DE DATOS – NO SE DEBE DILIGENCIAR</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> FUENTES DE DATOS</p> <p>¿Qué fuentes de datos se utilizan? (Mencione tablas de bases de datos internas y externas o métodos API). ¿De dónde se toman los datos? ¿Se pueden utilizar para realizar el objetivo del análisis?</p> <p>Opiniones de ciudadanos recogidas en procesos de participación territorial vinculados al seguimiento de políticas públicas.</p> <p>Datos proporcionados en los conjuntos compartidos en el curso (archivos con textos ya asociados a los Objetivos de Desarrollo Sostenible (ODS))</p> <p>Posibilidad de incluir fuentes externas como foros, encuestas en línea o reportes siempre que estén en lenguaje natural.</p>
---	--	---	--	---



SIMULACIÓN DE IMPACTO

¿Cuáles son los valores de costo/beneficio de las decisiones (in)correctas? ¿Cuáles son los criterios de éxito del modelo para su posterior despliegue? ¿Existen restricciones de equidad?

Una clasificación correcta (beneficio) generaría que las entidades asignen los recursos necesarios al ODS correcto, generando a su vez mayor impacto social, mejor percepción ciudadana y uso eficiente del presupuesto. Por el contrario, una clasificación incorrecta (costo), llevaría a invertir los recursos incorrectos a un ODS. Por



APRENDIZAJE (USO DEL MODELO)

¿El uso del modelo es por lotes o en tiempo real? ¿Con qué frecuencia se usa?

El uso del modelo es en tiempo real si los resultados son requeridos de manera inmediata, se puede usar por lotes para monitorear de manera rutinaria en espacios de tiempo largos. El modelo se usaría cada vez que se solicite una asociación entre una opinión y un ODS.



CONSTRUCCIÓN DE MODELOS

¿Cuántos modelos se necesitan? ¿Cuándo deben actualizarse? ¿De cuánto tiempo se dispone para generar el modelo (incluido el proceso de ingeniería de características y el análisis o evaluación del mismo)?

Se usaron cuatro modelos: Regresión Logística con TF-IDF, SVM lineal con TF-IDF por su buen desempeño en texto, un clasificador con embeddings para captar contexto semántico y un modelo de Árboles de Decisión que aporta interpretabilidad y reglas claras para entender cómo se clasifican las opiniones. Los modelos deben actualizarse periódicamente al disponer de nuevas



INGENIERÍA DE CARACTERÍSTICAS


¿Qué variables/características se utilizan en el modelo? ¿Qué agregaciones o transformaciones se aplican a las fuentes de datos originales – incluir las más importantes–?

Variables/características:

- Representaciones de texto como TF-IDF, frecuencia de palabras y embeddings
- Posible use de n-gramas para capturar contexto

Transformaciones:

- Limpieza de texto (eliminar stopwords, caracteres especiales, normalización)
- Conversión de frecuencias a representaciones numéricas (vectorización)
- Selección de las palabras más relevantes para la clasificación.

<p>ejemplo, se destinarían los recursos de educación cuando el ODS era el de pobreza. Esto genera desperdicio de recursos, desconfianza ciudadana y pérdida de oportunidades para resolver problemas urgentes.</p> <p>Un criterio de éxito para el despliegue del modelo son los validadores de precisión, recall y f1 score.</p> <p>Si hay restricciones de equidad, un ejemplo de esto es el lenguaje regional o variaciones idiomáticas, estos factores no deberían sesgar las predicciones (clasificaciones).</p>			<p>opiniones ciudadanas. El tiempo para generar cada modelo, incluyendo ingeniería de características y evaluación, conlleva semanas de trabajo dentro de las diferentes etapas del proyecto del curso.</p>	
	<div> <div>MONITOREO NO SE DEBE DILIGENCIAR</div>  </div>			

¿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?

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 when sharing.

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