|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TAREA DE APRENDIZAJE    Es un problema de aprendizaje supervisado. El modelo debe predecir a qué ODS (1: Fin de la pobreza, 3: Salud y bienestar, 4: Educación de calidad) corresponde una opinión ciudadana expresada en texto libre. Los resultados se observan en el momento en que se introduce una nueva opinión en la aplicación (tiempo real). | DECISIONES  ¿Cómo se convierten los resultados del modelo en recomendaciones o decisiones procesables para el usuario final?  Los resultados del modelo se convierten en una clasificación automática de las opinionesciudadanas. Esto permite a los responsables de políticas públicas:   * Detectar en qué temas (pobreza, salud o educación) hay más preocupaciones de la comunidad. * Priorizar recursos o acciones en función de la voz ciudadana. | PROPUESTA DE VALOR  ¿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?  El beneficiario final es el UNFPA y las instituciones públicas locales que diseñan políticas alineadas con los ODS. El problema específico que se aborda es la gran dificultad de analizar manualmente grandes volúmenes de texto. El riesgo es que un mal desempeño del modelo genere clasificaciones equivocadas que lleven a priorizar mal los recursos. | RECOLECCIÓN DE DATOS – NO SE DEBE DILIGENCIAR  ¿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? | FUENTES DE DATOS  ¿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?   * Opiniones ciudadanas recolectadas en procesos de participación comunitaria. * Datos textuales de encuestas abiertas o foros públicos asociados a planeación territorial. * Registros históricos de proyectos sociales con textos asociados a ODS. |
| 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?  Beneficio: reducir tiempo y costo en el análisis de grandes volúmenes de texto, facilitando decisiones basadas en evidencia.  Criterios de éxito: que el modelo tenga métricas de desempeño aceptables y que los decisores encuentren utilidad práctica en las clasificaciones.  Restricciones de equidad: asegurar que el modelo no tenga sesgos hacia ciertos grupos sociales o regiones. | APRENDIZAJE (USO DEL MODELO)  ¿El uso del modelo es por lotes o en tiempo real? ¿Con qué frecuencia se usa?  El modelo se usa en tiempo real, cada vez que se ingresa una nueva opinión en la aplicación web/móvil.  Puede complementarse con análisis por lotes cuando se procesen encuestas masivas.  La frecuencia de uso es continua, siempre que ciudadanos ingresen comentarios. |  | 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 requiere al menos un modelo de clasificación multiclase. El modelo debe actualizarse periódicamente (cada 6 meses o cuando se recolecten suficientes nuevos datos). El tiempo de construcción incluye: limpieza de datos, ingeniería de características, entrenamiento y validación. | 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: representaciones vectoriales del texto (Bag of Words, TF-IDF, embeddings).  Transformaciones:   * Normalización de texto (minúsculas, eliminación de stopwords, lematización). * Extracción de n-gramas para capturar expresiones frecuentes (“salud pública”, “acceso a educación”).   Características más importantes: frecuencia de palabras clave asociadas a cada ODS y vectores semánticos que capturan el contexto del texto. |
|  | MONITOREO NO SE DEBE DILIGENCIAR  ¿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](https://www.ownml.co/) when sharing. | [**OWNML.CO**](https://www.ownml.co/) |

# 