|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TAREA DE APRENDIZAJE    **Tipo:** Clasificación supervisada multiclase.  **Qué se predice:** la etiqueta ODS asociada a un texto en español que expresa una necesidad/observación ciudadana.  **Clases/Resultados posibles:** ODS1 (Fin de la pobreza), ODS3 (Salud y bienestar), ODS4 (Educación de calidad).  **Cuándo se observa el resultado:** inmediatamente al recibir el texto (inferencia en milisegundos); el modelo se entrena offline y se usa en batch para la entrega de la Etapa 1.  **Datos base:** 2.424 registros de entrenamiento (textos, labels), mapeando labels {1→ODS1, 3→ODS3, 4→ODS4}; 152 textos de prueba para etiquetar.  **Métrica principal:** F1 macro por desbalance de clases.  **Resultados** (CV 5-fold): Linear SVM (0.9724), Logistic Regression (0.9494), Naive Bayes (0.8034). Modelo seleccionado: Linear SVM (TF-IDF bigramas + acentos normalizados). | DECISIONES  **Objetivo operativo:** clasificar y **enrutar** cada texto al equipo responsable del ODS correspondiente (pobreza, salud, educación) para priorización y análisis.  **Reglas de acción sugeridas:**   1. **Asignar ODS** = predicción del modelo (SVM). 2. **Revisión manual** cuando el texto sea muy corto (<10 caracteres) o contenga solo números/URLs; (si se dispone del margen de decisión del SVM, revisar los percentiles más bajos). 3. **Tablero de priorización:** conteos por ODS, top términos por ODS, y lista de ejemplos “dudosos” para curaduría.   **Beneficio:** reduce tiempos de lectura manual y mejora la asignación de casos a programas/secretarías. | PROPUESTA DE VALOR  **Beneficiario:** área de planeación/social de una entidad local (alcaldía/universidad/ONG) que analiza voz ciudadana.  **Problema resuelto:** clasificación consistente y rápida de opiniones a ODS 1/3/4 para **priorizar** políticas y dirigir recursos.  **Valor:** +productividad analítica, métricas comparables en el tiempo, detección de tendencias por ODS.  **Riesgos y mitigación:**   * **Riesgo de sesgo/errores:** términos locales o jerga pueden producir confusiones; mitigar con revisión de “casos frontera” y ejemplos representativos. * **Riesgo de uso indebido:** el modelo **no** reemplaza criterio experto; incluir nota visible “asistente de apoyo, requiere verificación”. | 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  **Entrenamiento:** data/Datos\_proyecto.xlsx (Sheet1) con columnas **textos** (cadena) y **labels** (1/3/4).  **Prueba (entrega):** data/Datos de prueba\_proyecto.xlsx (Sheet1) con columna **Textos\_espanol**; se devuelve test\_etiquetado.xlsx añadiendo la columna **prediccion\_modelo**.  **Derivadas (features):** matriz TF-IDF con n-grams (1–2), normalización de acentos; sin stopwords embebidas para no perder señales específicas del dominio.  **Uso:** los datos son suficientes para cumplir el objetivo de la Etapa 1 (construcción y evaluación del modelo y etiquetado del test). |
| SIMULACIÓN DE IMPACTO  **Costos/beneficios (intuición):**  Acierto → texto bien enrutado al ODS → ahorro de tiempo y mejor focalización (beneficio alto).  Error → puede desviar la atención a un programa equivocado (costo moderado; mitigable con revisión de casos frontera).  **Criterios de éxito (para pasar a app en Etapa 2):** Macro-F1 ≥ 0.90 sostenido y F1 por clase ≥ 0.90.  Estabilidad del rendimiento en muestras de semanas distintas.  **Equidad:** revisar diferencias por longitud del texto y por presencia de términos regionales; si se detecta sesgo sistemático, ampliar datos o ajustar representación. | APRENDIZAJE (USO DEL MODELO)  **Modalidad de uso (Etapa 1):** batch sobre el Excel de prueba → se genera test\_etiquetado.xlsx.  **Frecuencia:** bajo demanda (cada corte de recolección).  Latencia de predicción: milisegundos por texto en CPU estándar (Python + scikit-learn).  **Etapa 2 (referencia del enunciado):** se prevé una API con /predict y, para rol experto, /retrain. |  | CONSTRUCCIÓN DE MODELOS  **Modelos evaluados:** Naive Bayes, Regresión Logística y **Linear SVM**. **Ganador:** SVM (**F1 macro 0.9724**) con TF-IDF (1–2).  **Actualización:** cada nuevo lote etiquetado (o mensual) si varía la distribución de textos; en Etapa 2 se comparará contra el baseline y se actualizará solo si mejora macro-F1.  **Tiempo de construcción:** preparación + entrenamiento < 2 min en CPU estándar para 2.4k textos; inferencia instantánea.  **Trazabilidad:** se guardan métricas CV, matriz de confusión y top-términos por clase para auditoría. (El Canvas hace parte de la Sección 1 del documento. | INGENIERÍA DE CARACTERÍSTICAS  **Variables de entrada:** texto libre en español.  **Transformaciones:**   * Normalización: minúsculas y **remoción de acentos**. * **TF-IDF** con **n-grams (1–2)** para capturar palabras y bi-gramas semánticos (p. ej., “salud pública”, “educación primaria”). * Opcional (no usado en baseline): lematización ligera y lista de stopwords en español si se valida que no degrade F1.   **Importancia de términos:** para SVM, se reportan coeficientes por clase (top-términos) como explicación cualitativa de cada ODS. |
|  | 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/) |

# 