

<div>TAREA DE APRENDIZAJE</div> <div></div> <div>Tipo de tarea: Aprendizaje supervisado de clasificación</div> <div>Predicción: Clasificar opiniones de ciudadanos de acuerdo a los términos ODS (tipo 1, tipo 3, y tipo 4) relacionados con "Fin de la pobreza", "Salud y bienestar" y "Educación de calidad"</div> <div>Resultados: La categorización de cada opinión en uno de los términos ODS</div> <div>Observar: El resultado se obtiene en segundos (dependiendo de la cantidad de datos)</div>	<div>DECISIONES</div> <div></div> <div>Las predicciones del modelo permiten generar recomendaciones que vinculan opiniones de la ciudadanía con políticas correspondientes a cada término ODS. Esto puede facilitar la toma de decisiones en la planificación de estrategias para atacar las iniciativas.</div>	<div>PROPUESTA DE VALOR</div> <div></div> <div>El beneficiario final es la UNFPA y las organizaciones que trabajan para lograr las metas de la Agenda 2030. El problema es que los datos proporcionados por los ciudadanos son muy grandes lo que causa que el análisis y clasificación manual requiere muchos recursos y expertos especializados. El modelo puede automatizar esta labor y reducir el tiempo de análisis considerablemente pero siempre habrá riesgo de que algunas opiniones no queden clasificadas correctamente.</div>	<div>RECOLECCIÓN DE DATOS</div> <div></div> <div></div>	<div>FUENTES DE DATOS</div> <div></div> <div>Los datos que se utilizaran en este proyecto vienen de opiniones textuales proporcionados por los ciudadanos respecto ODS 1,3, y 4.</div>
<div>SIMULACIÓN DE IMPACTO</div> <div></div> <div>La automatización del análisis textual reduce el tiempo de labor humano y ayuda a disminuir recursos ayudando a que se puedan destinar para otros propósitos. Sin embargo, el costo es que puedan haber clasificaciones erróneas por parte del modelo que pueden afectar la asignación de recursos públicos para este proyecto. Para medir el éxito del modelo, usaremos las métricas F1-score ya que refleja equilibrio entre precisión y recall.</div>	<div>APRENDIZAJE (USO DEL MODELO)</div> <div></div> <div>Si tenemos en cuenta que el ciudadano proporciona su respuesta y que esa sea su única función, podríamos decir que el usuario final en realidad es el funcionario y no el ciudadano. En este caso, el modelo utilizará procesamiento por lotes para poder entregar informes agregados. El modelo se puede ejecutar periódicamente ya que la clasificación en tiempo real de cada respuesta no es una prioridad dado que las decisiones del gobierno no se basan en un solo texto sino en tendencias agregadas de varios ciudadanos. El uso sería por medio de una aplicación web donde ciudadanos tendrían una ventana de tiempo (tal vez en una escala de tiempo semanal)</div>		<div>CONSTRUCCIÓN DE MODELOS</div> <div></div> <div>Este proyecto hará uso de 3 algoritmos diferentes en la fase experimental para clasificar texto: 1. Regresión logística 2. Máquina de soporte vectorial 3. Naive Bayes En la app final, solo uno será seleccionado haciendo una evaluación comparativa entre el mejor F1-score. Para generar el modelo toca procesar datos, vectorizar, entrenar, y comparar. Este tiempo está en escala de días pero en un escenario real no debería superar una escala de horas.</div>	<div>INGENIERÍA DE CARACTERÍSTICAS</div> <div></div> <div>Para poder analizar los textos toca realizar procesos de tokenización, lematización, y eliminación de stop words. Después de este preprocesamiento toca vectorizar el texto con TF-ID. Con esto podemos hacer el análisis de palabras más influyentes categorizadas por ODS.</div>
	<div>MONITOREO</div> <div></div>		<div></div>	