

Inteligencia de Negocios

Proyecto 1: Etapa 1

Grupo 7

Laura Julieth Carretero Serrano - 202214922

Juan David Rios Nisperuza - 202215787

Samuel Hernandez Diaz - 202213772

Sección 1: Entendimiento del negocio

Se busca desarrollar una descripción del problema u oportunidad de negocio, los objetivos y criterios de éxito, el impacto del proyecto en Colombia, y un enfoque analítico detallado (tipo de análisis, técnicas y algoritmos a utilizar).

Oportunidad /problema Negocio	<p>Actualmente, es fundamental procesar un gran volumen de información textual proporcionada por los ciudadanos, con el objetivo de identificar problemáticas sociales y proponer soluciones que se alineen con los Objetivos de Desarrollo Sostenible (ODS), en particular los ODS 3 (Salud y bienestar), 4 (Educación de calidad) y 5 (Igualdad de género). Dado que este tipo de análisis manual resulta costoso y requiere una gran inversión de recursos humanos, es necesario contar con expertos o sistemas automatizados que optimicen el proceso.</p> <p>Este proyecto se enfoca en la creación de una solución automatizada para la identificación y clasificación de opiniones vinculadas a los ODS. A través de este enfoque, no solo se atienden los objetivos empresariales, sino que también se incorpora un impacto social significativo, ya que el sistema permitirá escuchar, analizar y visibilizar las opiniones de las personas sobre temas críticos como el bienestar, la educación y la igualdad de género.</p> <p>Al automatizar este análisis, se mejorará la eficiencia en la identificación de problemas y la formulación de soluciones, especialmente en el contexto colombiano, donde se busca dar voz a las personas afectadas por diversas problemáticas sociales. La efectividad del proyecto dependerá de su capacidad para clasificar con precisión los textos según los ODS, lo que será clave para evaluar su éxito a largo plazo.</p>
Objetivos	<ul style="list-style-type: none">● Implementar una solución tecnológica que permita procesar y clasificar grandes volúmenes de opiniones y textos en función de los ODS, especialmente en temas de salud, educación e igualdad de género.● Ofrecer una herramienta que, además de cumplir objetivos empresariales, tenga un impacto social significativo al permitir que se escuchen y comprendan mejor las voces de las personas afectadas por problemáticas sociales críticas.● Desarrollar un sistema que sea aplicable y efectivo en el contexto local, facilitando la identificación de problemáticas relevantes para la sociedad colombiana.
Criterios de éxito	<ul style="list-style-type: none">● El éxito del proyecto dependerá de su capacidad para clasificar con alta precisión los textos según los ODS, asegurando que las opiniones y problemas se asignen correctamente a las categorías correspondientes.● Medir la efectividad de la automatización en términos de cuánto reduce el tiempo y los costos de procesamiento en comparación con el análisis manual tradicional.● Evaluar cómo la herramienta permite la identificación de problemas sociales que previamente no se estaban visibilizando, y cómo esto

	<p>conduce a soluciones más efectivas y rápidas.</p> <ul style="list-style-type: none"> • El grado de adopción y uso de la herramienta por parte de empresas, organizaciones no gubernamentales y entes gubernamentales en Colombia será un indicador clave de éxito. • La calidad del modelo para identificar correctamente las categorías de los ODS será clave para su éxito, evaluando la precisión y eficacia de su funcionamiento. • El sistema debe ser capaz de ajustarse y mejorar continuamente a medida que procesa más datos, mediante un proceso de reentrenamiento constante que optimice su desempeño a largo plazo.
Organización y rol dentro de ella que se beneficia con la oportunidad definida	<p>Entidades públicas de Colombia, organizaciones no gubernamentales y agencias internacionales (como el UNFPA) que trabajan en políticas de Desarrollo Sostenible. Los principales beneficiados son los analistas de datos y los responsables de la toma de decisiones, ya que la automatización del análisis de textos facilita sus tareas al permitirles procesar grandes volúmenes de información más rápidamente y tomar decisiones mejor fundamentadas sobre las políticas y estrategias a implementar.</p>
Impacto que puede tener en Colombia este proyecto.	<ul style="list-style-type: none"> • La automatización del análisis de texto permitirá a los gobiernos locales y nacionales identificar de manera más eficiente las necesidades sociales y responder a ellas con políticas públicas mejor informadas y alineadas con los Objetivos de Desarrollo Sostenible (ODS). • Al reducir el tiempo y los costos necesarios para analizar grandes volúmenes de opiniones ciudadanas, las entidades públicas y ONGs podrán priorizar y distribuir recursos de manera más efectiva, impactando positivamente en áreas clave como salud, educación y equidad de género. • El proyecto permitirá que las voces de los ciudadanos se escuchen de manera más clara y se tomen en cuenta en la toma de decisiones, promoviendo una mayor inclusión y representatividad en las políticas de desarrollo sostenible. • Al facilitar el análisis de datos y la identificación de problemáticas, el proyecto puede generar nuevas oportunidades para soluciones innovadoras y emprendimientos sociales, que aborden desafíos específicos relacionados con los ODS en el contexto colombiano. • Al enfocarse en objetivos como la salud, educación e igualdad de género, este proyecto puede contribuir significativamente a reducir las brechas sociales y mejorar la calidad de vida de las comunidades más vulnerables en Colombia.

<p>Enfoque analítico</p>	<p>Categoría de análisis: Analítica predictiva y descriptiva.</p> <p>Tipo de aprendizaje: Aprendizaje supervisado y no supervisado.</p> <p>Tarea de aprendizaje: Clasificación de texto, análisis de sentimientos y agrupamiento (clustering) de temas.</p> <p>Técnicas y algoritmos propuestos: Se ha elegido un enfoque analítico basado en aprendizaje supervisado, debido a la disponibilidad de etiquetas que definen los objetivos de desarrollo sostenible asociados a cada texto. La tarea de aprendizaje se centra en la clasificación, lo que facilitará la división de los datos en los tres objetivos de desarrollo sostenible asignados.</p> <p>Se utilizarán tres algoritmos:</p> <p>Decision Tree Classifier: Ofrece facilidad de implementación y capacidad de explicación, permitiendo identificar reglas de decisión claras para la clasificación.</p> <p>Random Forest Classifier: Un método basado en la combinación de múltiples árboles de decisión, que mejora la precisión y la robustez del modelo frente a datos no lineales y ruidosos.</p> <p>K-Nearest Neighbors (KNN): Un algoritmo de clasificación que clasifica los datos en función de la proximidad con sus vecinos más cercanos, ofreciendo un enfoque diferente y flexible para la identificación de patrones.</p> <p>La propuesta es útil para la empresa, ya que proporcionará un modelo con un alto porcentaje de precisión en la clasificación de opiniones en los objetivos deseados. Además, ofrecerá una visualización clara y comprensible de los resultados propuestos.</p>
---------------------------------	---

Sección 2 - 4:

La explicación completa de estas secciones se encuentra en el Notebook.

Sección 5: Mapa de actores relacionado con el producto de datos creado

En esta sección se busca describir organizaciones que podrían beneficiarse del modelo analítico, identificando actores clave dentro de la misma, sus beneficios y posibles riesgos.

Rol dentro de la empresa	Tipo de Actor	Beneficio	Riesgo
Analistas de Datos de UNFPA	Usuario-cliente	Los analistas de datos se benefician al tener acceso a herramientas y datos que facilitan el seguimiento y análisis de opiniones locales sobre temas relacionados con los Objetivos de Desarrollo Sostenible (ODS), permitiéndoles tomar decisiones más informadas.	Un riesgo potencial es la interpretación inadecuada de los datos, lo que podría llevar a decisiones erróneas que no reflejan la realidad de las necesidades locales.
Entidades Públicas	Financiador	Las entidades públicas que financian programas y proyectos relacionados con los ODS se benefician al mejorar la eficacia en la planificación y ejecución de dichas iniciativas, asegurando una inversión más estratégica de tiempo y dinero.	Existe el riesgo de que los recursos asignados sean desviados hacia proyectos que no estén alineados con los ODS, debido a intereses políticos u otros factores.
Organizaciones sin ánimo de lucro	Beneficiados	Las organizaciones del sector no lucrativo obtienen una mayor transparencia y eficiencia en la asignación de recursos y evaluación de proyectos, lo cual puede mejorar el impacto de sus iniciativas al utilizar los fondos proporcionados por el UNFPA de manera más efectiva.	Pueden enfrentar desafíos de sostenibilidad financiera a largo plazo, especialmente si los recursos no son constantes o están sujetos a restricciones.
Comunidades Locales	Beneficiados	Las comunidades locales, como principales beneficiarios, ganan mayor participación y voz en la toma de decisiones sobre su desarrollo y en el análisis de sus necesidades, fortaleciendo así la capacidad de influir en políticas y programas que les afectan directamente.	Puede existir el riesgo de que algunas comunidades no se comprometan totalmente o no participen activamente, lo que puede dificultar la implementación y el éxito de las iniciativas.

Fondo de las Naciones Unidas para la Infancia (UNICEF)	Asociado	La entidad podría utilizar los datos para ajustar sus programas de educación y protección infantil, adaptándolos a las necesidades específicas de las comunidades vulnerables.	Si los resultados del modelo no son precisos, los programas de protección infantil podrían ser insuficientes o mal dirigidos.
Dirección de Planeación Estratégica	Financiador	El uso de este modelo permite tomar decisiones estratégicas basadas en datos reales y alineados con los ODS, mejorando la asignación de recursos.	Si el modelo no cumple las expectativas, puede percibirse como una inversión ineficaz, afectando futuras decisiones de inversión tecnológica
Ciudadanos (sociedad civil)	Beneficiados	Sus opiniones son tomadas en cuenta en la planificación de políticas y proyectos alineados con los ODS, lo que podría mejorar sus condiciones de vida.	Si el modelo no interpreta correctamente las opiniones, los problemas reales de los ciudadanos podrían no ser atendidos de manera efectiva.
Alcaldías locales	Beneficiados	Pueden utilizar el modelo para diseñar políticas más específicas en educación y equidad de género (ODS 4 y 5) adaptadas a las necesidades de cada comunidad.	Si el modelo no captura bien las particularidades locales, las políticas podrían no tener el impacto deseado, desperdiciando recursos municipales.
Organización Internacional del Trabajo (OIT)	Asociado	Aprovecha el modelo para alinear sus programas de desarrollo de competencias laborales (ODS 4) con las necesidades percibidas por los ciudadanos en las áreas de formación y empleo.	Un mal rendimiento del modelo podría llevar a la implementación de programas laborales no alineados con las verdaderas demandas del mercado y la población.

Sección 6: Trabajo en equipo

Este entregable busca que el equipo detalle las actividades realizadas por cada miembro, especificando los roles, tareas, y algoritmos en los que trabajaron, así como el tiempo invertido en cada actividad. Además, se deben describir los retos enfrentados durante el proyecto y las soluciones planteadas para resolverlos. Con base en la planificación inicial y la experiencia durante el desarrollo del proyecto, se busca reflexionar sobre la distribución de 100 puntos entre los integrantes, evaluando el aporte individual, e identificar oportunidades de mejora para futuras entregas.

Integrante	Rol	Tarea Realizada	Tiempo dedicado	Algoritmo trabajado	Puntos repartidos
Samuel	Lider de Proyecto y líder de analítica	Trabajó en la Sección 4 (Resultados), que incluye la descripción de los resultados obtenidos, el análisis de las métricas de calidad, y la estrategia basada en los resultados de los modelos analíticos. También trabajó en la Sección 3 (Modelado y evaluación), aplicando diferentes algoritmos de aprendizaje automático.	6.5	Decision Tree Classifier	33.33
Laura	Lider de negocio	Trabajó en la Sección 1 (Entendimiento del negocio y enfoque analítico), que involucra la identificación de oportunidades de negocio, objetivos, criterios de éxito y el enfoque analítico. También contribuyó a la Sección 3 (Modelado y evaluación) y trabajó en la Sección 5 (Mapa de actores), donde describió a las organizaciones que se pueden beneficiar del modelo analítico creado.	5	Random Forest Classifier	33.33
Juan	Lider de datos	Encargado de la Sección 2 (Entendimiento y preparación de los datos),	5	K-Nearest Neighbors (KNN)	33.33

		que implica el perfilamiento, análisis de calidad y preparación de datos según las técnicas y algoritmos seleccionados. También colaboró en la Sección 3 (Modelado y evaluación).			
--	--	---	--	--	--

Retos enfrentados en el proyecto:

Uno de los principales retos que enfrentamos fue la coordinación del trabajo en equipo. Cada integrante tenía responsabilidades específicas, pero al ser un proyecto que requería el manejo de datos, desarrollo de modelos analíticos y la comprensión del negocio, hubo momentos en que la comunicación se volvió un desafío. Esto se sumó al reto de integrar nuestras diferentes áreas de trabajo, mientras algunos tenían más experiencia en la parte técnica, otros destacaban en la comprensión del negocio.

Otro desafío fue elegir el algoritmo adecuado. Tanto el segundo como el tercer algoritmo presentaron métricas muy similares, lo que hizo difícil decidir cuál utilizar. A pesar de que ambos tenían resultados comparables, era crucial analizar en detalle sus características y el comportamiento en distintos escenarios para tomar la mejor decisión.

Formas planteadas para resolverlo:

Para abordar los problemas de coordinación y comunicación, decidimos tener reuniones de sincronización más frecuentes y definir de forma más clara las responsabilidades de cada miembro en cada fase del proyecto. Esto nos ayudó a mantenernos alineados y a asegurar que todos entendían el progreso y los próximos pasos.

Decidimos optar por el algoritmo que mostró métricas más equilibradas, que en este caso fue el KNN (K-Nearest Neighbors). Aunque ambos algoritmos presentaron resultados similares, el KNN destacó por mantener un balance consistente en sus métricas, lo que indicaba una mayor estabilidad y previsibilidad en su desempeño. Esta elección se basó en la necesidad de un algoritmo que no solo tuviera buenos resultados en promedio, sino que también fuera fiable en una variedad de situaciones, asegurando así una mejor adaptabilidad y precisión en las predicciones.

Reflexión sobre la repartición de los puntos:

En cuanto a la repartición de los puntos, consideramos que la contribución de cada miembro del equipo fue equivalente y esencial para el éxito del proyecto. Aunque cada integrante tuvo roles diferentes y abordó tareas específicas, todos aportaron de manera significativa, complementando las habilidades y conocimientos de los demás.

El enfoque del equipo se basó en la colaboración continua entre los miembros. Todos participaron activamente en las discusiones y decisiones clave, y cada miembro aportó desde su área para complementar las habilidades de los demás. Esto permitió que el proyecto avanzara de manera fluida y que cada etapa se completara con el nivel de detalle y calidad requerido.

Los puntos se repartieron de manera equilibrada porque cada integrante, sin importar su rol, asumió tareas importantes que contribuyeron al resultado final. Desde la definición del enfoque estratégico hasta la preparación de los datos, el modelado, y la interpretación de los resultados, todos los esfuerzos fueron cruciales para lograr un producto final robusto y alineado con los objetivos del negocio.