

Proyecto 1 – Etapa 2: Analítica de Textos

**Universidad de los Andes Departamento de Ingeniería de Sistemas y
Computación
ISIS 3301 Inteligencia de Negocios
202320**

**Andrés Francisco Borda 201729184
Juan Pablo Lora Hernández 202012524
Gabriela Vargas Rojas 202013830**

Contenido

1.1	Proceso de automatización del proceso de preparación de datos, construcción del modelo, persistencia del modelo y acceso por medio de API:...	2
1.2	Desarrollo de la aplicación y justificación.....	2
1.3	Resultados	6
1.4	Trabajo en Equipo	6

1.1 Proceso de automatización del proceso de preparación de datos, construcción del modelo, persistencia del modelo y acceso por medio de API:

Para realizar la automatización del proceso de preparación de datos y la construcción de este modelo, se implementó un pipeline en el cual se adecuaron las funciones de preprocesamiento que se utilizaron en la primer etapa del proyecto. Así mismo, el resultado del pipeline fue exportado como un .joblib para que la organización pueda tener una persistencia de este modelo sin la necesidad de tener que ejecutarlo nuevamente.

Para el acceso se utilizó un API llamada FastAPI, la cual permite un excelente desarrollo para las páginas web. Está diseñada para proporcionar un servicio muy veloz y una fácil conexión en este caso con el front y la carga del modelo de analítica de textos. De esta manera, se utilizó para cargar el archivo del pipeline y que este pudiera ser utilizado en la aplicación web de una manera muy rápida.

FastAPI así mismo, posee una documentación bastante sencilla y clara lo que facilita cualquier tipo de entendimiento o aprendizaje que se quiera hacer sobre el servicio.

De esta manera, se hizo el respectivo junte entre el pipeline del modelo analítico y los servicios API a través de un sencillo programa en Python. Todo esto, con el fin de desarrollar la aplicación web para que el usuario interactúe con los resultados obtenidos y la necesidad de procesamiento de nuevos datos para los ODS trabajados.

1.2 Desarrollo de la aplicación y justificación

Para la consolidación del desarrollo de la aplicación, se tuvieron en cuenta diferentes tipos de usuario que podrían hacer uso de la aplicación, tomadas del mapa de actores que se realizó en la etapa 1:

Organización/Apartado	Rol dentro de la empresa	Tipo de Actor	Beneficio	Riesgo
Equipo de Investigación de la Universidad de los Andes	Investigadores	Usuarios Internos	Los investigadores de la Universidad de los Andes utilizan el producto de datos para realizar investigaciones académicas relacionadas con los ODS en Colombia, lo que contribuye al avance del conocimiento en el campo del desarrollo sostenible.	Si los datos no son precisos o completos, las investigaciones académicas basadas en ellos podrían carecer de fiabilidad.
Estudiantes de la Universidad de los Andes	Participantes	Usuarios Internos	Los estudiantes pueden aprender sobre la aplicación práctica de los ODS en Colombia a través de proyectos y análisis basados en el producto de datos, lo que enriquece su	Si la participación de los estudiantes en la recopilación o análisis de datos no se gestiona adecuadamente, podría haber desafíos logísticos o éticos.

			educación y conciencia sobre el desarrollo sostenible.	
UNFPA Colombia	Socio Estratégico	Beneficiarios/Colaboradores	UNFPA Colombia se beneficia al contar con un producto de datos que le permite evaluar la percepción de las comunidades locales en relación con los ODS en Colombia. Esto facilita la toma de decisiones informadas y la planificación de programas más efectivos.	Si los datos recopilados y analizados no reflejan con precisión la realidad de las comunidades, esto podría llevar a decisiones ineficaces o inadecuadas.
Comunidades y Grupos Locales en Colombia	Beneficiarios/Interesados	Beneficiarios	Las comunidades locales se benefician al tener sus voces y necesidades representadas en el proceso de planificación de proyectos y	Si las comunidades no confían en el proceso de recopilación de datos o no comprenden su importancia, podrían no

			programas relacionados con los ODS en Colombia.	participar de manera efectiva.
--	--	--	-------------------------------------------------	--------------------------------

De esta manera, e intentando seguir el objetivo principal propuesto, se determinó que la aplicación debía estar desarrollada principalmente para una persona perteneciente a la UNFPA Colombia, donde pudiera hacer uso directamente del modelo teniendo la posibilidad de ingresar nuevas cadenas de texto (testimonios de personas) y poder predecir a que ODS pertenece con el modelo analítico desarrollado sobre el algoritmo de Regresión Logística.

En este caso, la conexión con la oportunidad principal para este negocio será poder hacer un análisis automatizado de las voces de los habitantes locales sobre las problemáticas de su entorno particular y así llevarlas de la mano al incluirlas en la aplicación de los objetivos de desarrollo sostenible, permitiendo de esta forma que el enfoque de recursos y tiempo que era utilizado para el análisis de estos datos, sea más bien utilizado para directamente llevar esos testimonios hacia su uso principal dentro de la agenda propuesta para el cumplimiento de los Objetivos de desarrollo sostenible.

Finalmente, este desarrollo y conclusión fue alcanzado gracias al trabajo interdisciplinar que se realizó con el grupo de estadística, donde después de algunas reuniones y la muestra de la aplicación web con el modelo incluido, se determinó que la mejor forma de cumplir el objetivo propuesto era atacar el proceso de negocio de automatización que se presentaba en el proyecto. Este trabajo fue importante, ya que, permitió tomar una perspectiva de trabajo más enfocada hacia un usuario final y el uso del modelo analítico directamente para este, de una forma sencilla y amigable pero útil así mismo para el entorno trabajado. Adicionalmente, permitió a este equipo de trabajo presentar los resultados de una manera mucho más centrada en cualquier tipo de público e interactiva en el hecho de la facilidad de uso de la aplicación final.

La persona del grupo de estadística nos aportó su conocimiento al respecto de cómo se concluye generalmente sobre el enfoque analítico construido y de qué manera se puede incluir el desarrollo del trabajo proporcionando una alta accesibilidad sin entrar en la visualización directa de las métricas matemáticas obtenidas.

1.3 Resultados

Los resultados obtenidos serán mostrados en el vídeo subido a la plataforma de YouTube, donde se muestran las posibilidades que ofrece la aplicación al usuario específicamente la manera en la que se pueden proporcionar nuevos datos escribiendo en un cuadro de texto o subiendo un archivo de testimonios de persona. Estos datos, serán procesados por nuestro modelo y se le mostrará al usuario el resultado de la predicción que obtiene el mismo.

La aplicación final, fue establecida bajo las directrices de una facilidad de uso y una interfaz muy sencilla para el usuario con el fin de que pueda aportar al problema de negocio sin mayores complicaciones. Así mismo, se le otorga al usuario final todo el modelo construido de manera comprimida en el pipeline que se mencionó anteriormente, esto con el fin de darle una persistencia y de qué si es necesario se pueda analizar o replicar todos los pasos de preprocesamiento que se realizaron sobre los datos y cómo esto favoreció en las métricas del modelo.

1.4 Trabajo en Equipo

Estudiantes	Rol	Horas trabajadas	Tareas hechas	Desafíos	Puntaje
Andrés Borda	Líder de proyecto – Ingeniero de Datos	10	Automatización para el Pipeline – Documento – Vídeo Resultados	Investigación para la automatización	33,33
Juan Pablo Lora	Ingeniero de software responsable de diseño y resultados	10	Diseño de la aplicación web – Documento – Vídeo Resultados	Investigación para el uso de API	33,33
Gabriela Vargas	Ingeniero de Software responsable de la aplicación final	10	Implementación de la Aplicación Final – Documento – Vídeo Resultados	Investigación de la usabilidad del usuario final.	33,33