

Proyecto - Entrega 2 - Grupo 11

▪ (20%) Proceso de automatización del proceso de preparación de datos, construcción del modelo, persistencia del modelo y acceso por medio de API:

Descripción del proceso e implementación realizados por el ingeniero de datos, acompañado del código. En este punto es importante que revise si es posible enriquecer el modelo analítico construido previamente con integración de nuevas columnas del mismo archivo original o de otras fuentes para obtener mejores resultados tanto desde el punto de vista cuantitativo, como cualitativo. Se sugiere utilizar pipelines para este proceso y un Framework para desarrollo de la API.

El proceso realizado por el ingeniero de datos consiste en la preparación del modelo que se va a usar para el procesamiento de texto, también realiza el proceso de creación del pipeline para poder exportar el modelo para ser usado posteriormente en el desarrollo de la API, aquí también se incluye el proceso de vectorización por medio de TFIDF. No se consideró necesario un cambio en el modelo creado anteriormente debido a que las métricas nos resultaron muy confiables y no encontramos manera de mejorarlas.

▪ (40%) Desarrollo de la aplicación y justificación.

Descripción del usuario/rol de la organización que va a utilizar la aplicación, la conexión entre esa aplicación y el proceso de negocio que va a apoyar (si aplica), y la importancia que tiene para ese rol la existencia de esta aplicación. Adicionalmente, el desarrollo de una aplicación web o móvil para interactuar con el resultado del modelo (y su calidad) a partir de un texto o textos datos por el usuario.

El usuario de la organización que va a utilizar la aplicación es un operador que es responsable de verificar los mensajes que el modelo arroja como suicidas para analizarlos individualmente y continuar con el protocolo del caso. Esta aplicación va a apoyar a este proceso debido a que el operador no va a tener que revisar cada mensaje, sino que va a poder tener una referencia de los mensajes que tienen mucha probabilidad de tener contenido suicida, por lo que la identificación de los mismos será mucho más rápida.

▪ (18%) Resultados.

Video de máximo ocho (7) minutos con la descripción y visualización en la aplicación de los resultados del modelo, que permita a un rol dentro de la organización comprenderlos y usarlos. El video debe simular la interacción del usuario final con la aplicación y describir dos acciones que puede realizar como resultado de dicha interacción, haciendo énfasis en la forma como el resultado del modelo aporta en esas acciones.

▪ (10%) Trabajo en equipo.

Para el trabajo en equipo se sugieren los siguientes roles para cada integrante. Un integrante puede desempeñar más de un rol.

- *Líder de proyecto: Daniel Felipe Martínez. Define las tareas que cada integrante debe realizar.*

- *Ingeniero de datos: Daniel Felipe Martínez. Hizo el proceso de creación del pipeline teniendo en cuenta el modelo generado en la etapa 1 del proyecto. Hace el proceso de desarrollo del API para el despliegue del modelo. Describe el proceso de automatización del proceso de preparación de datos. Hace la justificación para el desarrollo de la aplicación.*

Número de horas: 5

El tiempo fue principalmente por la demora que se tomó un error en el momento de importar el modelo para ser usado en el API. Se probaron muchas posibles soluciones y ninguna logró dar un buen resultado para poder continuar con el despliegue del API.

- *Ingeniero de software: Santiago*

- *Ingeniero de software: Santiago*

Puntaje

Daniel Felipe Martínez : 75

Santiago Tangarife: 25

Nicole: 0