

Documento Proyecto 1 Etapa 2 - DataSquad
Datos Suicidio - Salud mental

Diego Camelo (202011662)
Nicolás Rodríguez (201822980)
Isabela Ortiz (201631748)

Tabla de contenido

1. Automatización del proceso
2. API
3. Desarrollo de la aplicación
4. Trabajo en equipo
5. Referencias

1. Automatización del proceso

Para llevar a cabo la automatización del proceso de preparación de los datos y la creación del modelo, se creó un pipeline. Este pipeline utiliza el `TfidfVectorizer` como transformador para preparar los datos en formato de texto a número, que es la manera como después se puede aplicar el algoritmo de clasificación. Para la creación del pipeline intentamos crear los transformadores personalizados para ir preparando los datos paso a paso como en la etapa 1 del proyecto, sin embargo, nos dimos cuenta que con el transformador `TfidfVectorizer` ya se hace el preparamiento necesario, por lo cual se dejó este solamente.

El pipeline crea luego un modelo de redes neuronales. Un algoritmo de redes neuronales es un modelo matemático evolutivo conformado por “neuronas” que reciben entradas y generan salidas con un peso específico para aprender de sí mismos y enseñar la solución más óptima a un problema. Su mejor desempeño se da en tareas de detección de características y soluciones. Para este caso se consideró que era un algoritmo óptimo dada su habilidad para detectar características y dar resultados con una alta precisión.

Para la selección de hiperparámetros, se decidió probar varias combinaciones entre la cantidad de neuronas en la capa escondida y la cantidad de iteraciones de aprendizaje de las neuronas. Después de hacer varias pruebas en la etapa 1 del proyecto, se decidió crear un modelo con un total de 10 neuronas iterando 500 veces, ya que de esta manera conseguimos puntajes $f1$ más que satisfactorios (0.91).

Los hiperparámetros utilizados fueron:

```
hidden_layer_sizes=(10,10,10), max_iter=500, alpha=0.0001, solver='adam',  
random_state=21, tol=0.000000001
```

El modelo creado por el pipeline se guarda en un archivo `.joblib` utilizando `dump` para que luego pueda ser cargado y utilizado.

2. API

El desarrollo del api se realizó utilizando el framework “Fastapi”. Se desarrolló de tal manera que se recibe un json con un mensaje y es procesado para convertirlo en un String, que luego se convierte en una serie, para aplicar el modelo persistente que se obtuvo del pipeline. El modelo arroja 0 si la etiqueta no es suicida y 1 si la etiqueta sí indica que puede haber tendencias suicidas.

3. Desarrollo de la aplicación

Una organización que puede estar interesada en el proyecto es una entidad del sector de la salud, como por ejemplo, el Ministerio de Salud. Esta entidad puede estar interesada en disminuir la cantidad de suicidios anuales y tomar medidas para mejorar la salud mental del país. Concretamente, en el Ministerio de Salud existe un grupo de Gestión Integrada para la Salud Mental (Organigrama Ministerio de salud). Este grupo busca capacitar a personas y crear lineamientos con el fin de garantizar el derecho a la salud mental en Colombia. Por esto, puede ser de utilidad para este grupo promover la aplicación que proponemos en las capacitaciones de tal manera que por ejemplo médicos psiquiatras o psicólogos puedan utilizar la aplicación para apoyar procesos con pacientes que puedan estar teniendo pensamientos suicidas.

Para esto se desarrolló la aplicación, de tal manera que el médico o psicólogo pueda ingresar texto que el paciente esté publicando en redes sociales, o escribiendo en otros medios para determinar si puede o no estar considerando el suicidio y tomar medidas preventivas. De esta manera, la aplicación busca apoyar el diagnóstico y acompañamiento que los profesionales de la salud mental le brindan a sus pacientes.

La aplicación creada cuenta con un cuadro para ingresar un texto. Al oprimir el botón de calcular, se le aplica el modelo y sale un mensaje con el resultado, que indica si la persona podría presentar tendencias suicidas o no.

A continuación se muestra la aplicación desarrollada:

Evaluador de salud mental

Evaluar Mensaje:

La persona podría presentar tendencias suicidas

4. Trabajo en equipo:

Roles

Líder del proyecto - Isabela Ortiz

Ingeniero de datos - Isabela Ortiz

Ingeniero de software responsable del diseño de la aplicación y resultados: -
Nicolas Rodriguez

Ingeniero de software responsable de desarrollar la aplicación final - Diego
Camelo

Horas trabajadas por cada integrante:

Isabela Ortiz: 5 h

Nicolás Rodriguez: 3 h

Diego Camelo: 5 h

Tareas desarrolladas:

Como grupo tuvimos una primera reunión para definir qué roles iba a tomar cada integrante y entender de qué se trataba el proyecto. En esta reunión nos dividimos el trabajo de la siguiente manera: Nicolás se encargaría de la parte de la automatización y desarrollo del pipeline, Diego se encargaría de la conexión del pipeline con el API e Isabela se encargaría de crear la aplicación.

Tuvimos algunos problemas intentando hacer el pipeline con transformaciones personalizadas por lo cual Isabela tomó el liderazgo para desarrollar el pipeline y lo hizo con el TfidfVectorizer que no necesita transformaciones personalizadas. Luego, Diego tomó el liderazgo en conectar el API con el pipeline. Tanto Diego como Isabela diseñaron la manera en la que la aplicación funciona y a qué usuario le podría servir. Diego se encargó de crear el front de la aplicación. Por último, tanto Diego como Isabela registraron en video los resultados e hicieron el documento.

Puntos para repartir:

Isabela Ortiz: 40 puntos

Nicolás Rodriguez: 20 puntos

Diego Camelo: 40 puntos

Para mejorar:

Tuvimos inconvenientes para el desarrollo del proyecto como equipo debido a que no lo empezamos con la suficiente anterioridad y uno de los integrantes tuvo problemas durante el fin de semana para poder contribuir al proyecto. Por esta razón, debemos empezar a trabajar con una mayor anterioridad para que así cada integrante pueda organizar su tiempo y dedicarle el tiempo necesario al proyecto.

Link del repositorio de GitHub:

https://github.com/dcamelog/BIP1_2

Link del video de Youtube

<https://youtu.be/qqMjTIQATsQ>

5. Referencias

Ministerio de Salud. (2021). Gestión Integrada para la salud mental.

Recuperado de:

<https://www.minsalud.gov.co/salud/publica/SMental/Paginas/gestion-integrada-para-la-salud-mental.aspx>

Ministerio de Salud. (2022). Organigrama. Recuperado de:

<https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/SG/GT/organigrama-minsalud.pdf>