

DEPARTAMENTO DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

**Inteligencia de negocios 202220 – Laboratorio 4**

PROFESORA: Haydemar Nuñez

|  |  |  |  |
| --- | --- | --- | --- |
| **Nombres** | **Apellidos** | **Código** | **Login** |
| María Sofía | Álvarez López | 201729031 | ms.alvarezl |
| Brenda Catalina | Barahona Pinilla | 201812721 | bc.barahona |
| Alvaro Daniel | Plata Márquez | 201820098 | ad.plata |

**Informe de laboratorio #4**

**Escenarios de prueba para el API y análisis:**

Para probar los escenarios, puede encontrar cada uno de los Json en la carpeta “Test” . En los escenarios se pondrán los pantallazos de la prueba y el resultado que dio. Además de puntualizar el nombre de la prueba (del archivo)

* **Predicción coherente:**

Nombre del archivo: R2Coherente1

**Interfaz de usuario gráfica, Texto

Descripción generada automáticamenteInterfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente**

Para este escenario calculamos en Jupyter el y este es el mismo que el que obtuvimos con la API.

**Nombre del archivo: PrediccionCoherente1**

**Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente con confianza mediaInterfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente**

En esta predicción se puede ver que se ponen a prueba dos casos, y podemos ver que el life expectancy que resulta es muy cercano al que es en realidad. Para el primer caso, el valor de la variable de interés es de 74.94, cuando en realidad debió ser de 73.7. Para la segunda, la predicción dio 61.73, siendo en realizada 59.2.

* **Predicción incoherente**
* **Hay fallas**

Nombre del archivo: FallaUnnamed

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente Texto

Descripción generada automáticamente

Podemos ver que ocurre un error, al realizar una prueba sin la columna “Unnamed”. Esto ocurre porque en el modelo teníamos en los datos esta columna, que, en el pipeline eliminamos para hacer el modelo final de predicción. Con lo anterior generamos un problema ya que requerimos de esta columna para que el pipeline y por consiguiente el modelo funcione.

**Estrategia por desarrollar sobre el software para mitigar incoherencias en el resultado y fallas en el sistema:**

Una posible estrategia es mejorar el pipeline, en esta podríamos tener en cuenta la posibilidad de que no sea agregado la columna “Unnamed”, por lo que, el modelo debería correr exitosamente con o sin esta columna.

Por otro lado, podríamos dar un mejor manejo a los valores nulos, nosotros los tratamos con la media, pero se podría probar alternativas como la eliminación de estos datos o la imputación con otro valor significativo que no sea la media.

**Bono:**

1. Construir transformaciones personalizadas e incluirlas en el pipeline y garantizar que el proceso completo es correcto.

El pipeline que usamos para la realización de este laboratorio

2. Desplegar la API en un servidor gratuito como Heroku para que pueda prestar servicio a cualquiera haciendo uso de una URL.

3. Implementar la estrategia para mitigación de errores identificados en los escenarios y documentados por ustedes en el documento de entrega.