**UNIVERSIDAD DE LOS ANDES DEPARTAMENTO DE INGENIERIA DE SISTEMAS Y COMPUTACIÓN**

****

**PROYECTO 1**

**ETAPA 2**

**INTELIGENCIA DE NEGOCIOS**

**FELIPE JAIMES - 202213793**

**MARCOS ESPAÑA - 202124714**

**JAIRO FIERRO - 202226326**

**2025-10**

Contenido

[1.Introducción 3](#_Toc194133944)

[2.Sección 1 3](#_Toc194133945)

[2.1 Primer Endpoint 3](#_Toc194133946)

[2.2 Segundo Endpoint 3](#_Toc194133947)

[2.2.1 Definiciones de reentrenamiento 3](#_Toc194133948)

[2.2.1.1 Reentrenamiento total (batch) 3](#_Toc194133949)

[2.2.1.2 Reentrenamiento incremental (online learning) 3](#_Toc194133950)

[2.2.1.3 Reentrenamiento periódico por validación 4](#_Toc194133951)

[3.Sección 2 4](#_Toc194133952)

[3.1 Descripción del rol del usuario 4](#_Toc194133953)

[3.2 Preguntas respecto a la aplicación 4](#_Toc194133954)

### 1.Introducción

### 2.Sección 1

### 2.1 Primer Endpoint

Este endpoint recibe una o mas entradas de noticias con titulo y descripción para poder hacer una predicción donde se dice si la noticia es falsa o no y también se dice la probabilidad de la predicción.

### 2.2 Segundo Endpoint

### 2.2.1 Definiciones de reentrenamiento

### 2.2.1.1 Reentrenamiento total (batch)

* Descripción: Se entrena un nuevo modelo desde cero utilizando la totalidad del dataset disponible, es decir, la combinación de los datos históricos (viejos) y los datos recién adquiridos (nuevos).
* Ventaja: Se asegura que el modelo no olvida los datos del pasado y captura patrones globales de todo el conjunto de datos, ofreciendo un aprendizaje completo y consistente.
* Desventaja: Costoso en términos de tiempo y recursos computacionales, especialmente cuando el volumen de datos crece significativamente, lo que puede hacerlo impracticable en sistemas con datasets masivos.

### 2.2.1.2 Reentrenamiento incremental (online learning)

* Descripción: Se actualiza el modelo existente incorporando únicamente los nuevos datos a medida que llegan, sin necesidad de reentrenar desde cero con el dataset completo.
* Ventaja: Muy eficiente y rápido, ya que evita procesar nuevamente los datos antiguos. Es ideal para escenarios con datos en streaming o sistemas que requieren adaptarse en tiempo real.
* Desventaja: No todos los modelos o algoritmos soportan aprendizaje incremental de forma nativa, y puede haber riesgo de olvidar patrones importantes de los datos históricos (catastrophic forgetting) si no se implementa cuidadosamente.

### 2.2.1.3 Reentrenamiento periódico por validación

* Descripción: Se reentrena el modelo de forma programada (según un cronograma fijo o cuando una métrica de rendimiento cae por debajo de un umbral predefinido), evaluando previamente si es necesario en función de su desempeño actual.
* Ventaja: Reduce el reentrenamiento innecesario al realizarlo solo cuando hay evidencia de deterioro en el modelo, optimizando recursos y tiempo.
* Desventaja: Existe el riesgo de dejar pasar datos nuevos relevantes que podrían mejorar el modelo antes de que se detecte una caída en el rendimiento, lo que podría retrasar su actualización óptima.

### 3.Sección 2

### 3.1 Descripción del rol del usuario

### 3.2 Preguntas respecto a la aplicación