



Maestría
en Inteligencia
Analítica de Datos



Proyecto - Entrega 2

Descripción general

El objetivo de este proyecto es desarrollar y desplegar un producto de analítica de datos. Cada equipo debe seleccionar un problema y un conjunto de datos asociado. A lo largo del semestre deben ir desarrollando la solución, desde la comprensión del problema de negocio y los datos, hasta el despliegue de la solución analítica. El nivel esperado de desarrollo de este producto es de un **prototipo funcional**. La solución debe consistir de **modelos supervisados empaquetados**, una **API** para su servir las inferencias y un **tablero** que permita emplear el modelo a través de la API y visualizar otros datos relevantes para el usuario.

Avance esperado

Terminada la primera entrega, para la segunda entrega se espera que cada equipo haya desarrollado las siguientes actividades:

Semana 4

En esta semana el equipo debe por lo menos:

- Desarrollar las primeras versiones de los modelos.
- Realizar experimentos con los modelos empleando MLflow para versionar los modelos y los resultados de sus experimentos.
- Desarrollar el tablero de acuerdo con la maqueta diseñada.
- Documentar.

Semana 5

En esta semana el equipo debe por lo menos:

- Desarrollar nuevas versiones de los modelos, comparar y seleccionar mejores alternativas.



- Realizar experimentos con los modelos empleando MLflow para versionar los modelos y los resultados de sus experimentos.
- Desarrollar el tablero de acuerdo con la maqueta diseñada.
- Documentar.

Entrega 2

Al final de la semana 5 (**domingo 9 de noviembre**) deberán entregar un reporte y soportes sobre los siguientes elementos:

- Modelos desarrollados.
- Experimentos realizados y su soporte en MLflow.
- Tablero desarrollado de acuerdo con la maqueta.
- Repositorio Git con todo el código desarrollado. Aquí deben reflejarse los aportes de todos los miembros a través de commits.
- Reporte de trabajo en equipo.

Entregables

- Reporte de máximo 10 páginas que incluya:
 - Resumen sobre el problema que están abordando (máx. 1 página):
 - Contexto del problema.
 - Pregunta de negocio y alcance del proyecto.
 - Breve descripción de conjuntos de datos a emplear.

Resalte posibles cambios con respecto a la primera entrega

 - Modelos desarrollados y su evaluación.
 - Observaciones y conclusiones sobre los modelos.
 - Descripción del tablero desarrollado y la funcionalidad que éste ofrece.

- Repositorios con todo el código (accesibles y con evidencia sobre el uso de repositorios por parte de los miembros del equipo).
- Fuentes de los modelos desarrollados.
- Fuentes del tablero desarrollado.
- Pantallazos de experimentos registrados en MLflow en una máquina de AWS EC2 (debe ser visible el usuario e IP de la máquina en EC2, y la IP en MLflow). *Mantenga su máquina en EC2 con*



MLflow detenida (no la termine).

- Reporte de trabajo en equipo de máximo 1 página.

Nota: los soportes son parte fundamental de la entrega. Su no entrega lleva a una alta penalización.

Nota 2: si bien el trabajo es en equipo, la nota es individual, luego es necesario que cada miembro del equipo demuestre su contribución al proyecto, tanto a través de los **commits en el repositorio**, como a través del **reporte** de trabajo en equipo.

Recomendaciones

1. El objetivo del proyecto es lograr un buen producto, bien soportado y claramente desarrollado. Justifique adecuadamente sus decisiones, observaciones y conclusiones.
2. Sea conciso y eficiente con el espacio. En un buen reporte cada gráfica y afirmación importa.
3. Es un trabajo en equipo. Defina los ítems de trabajo, asígnelos entre los miembros del equipo, defina fechas de entrega y revisión interna. Discuta los resultados, observaciones y conclusiones. Priorice tareas y resultados a incluir.
4. Empiece a trabajar prontamente y discuta con el equipo de tutores su avance y resultados.