

Micro-proyecto - Entrega 2

Descripción general

El objetivo de este proyecto es desarrollar y desplegar un producto basado en 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 22 de febrero**) 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.

Nota 3: el reporte debe ser de **máximo 10 páginas**. Si se entrega un reporte de mayor extensión, solo se calificarán las primeras 10 páginas.

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, asígneles 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.

Rúbrica

A continuación se presenta la rúbrica de calificación, con un resumen de cada punto considerado. La descripción completa de cada instrucción la puede encontrar en el numeral asociado.

Entregable	Rúbrica
Resumen sobre el problema.	<p>10 puntos</p> <ul style="list-style-type: none"> Se identifica adecuadamente el problema, su contexto, pregunta de negocio, alcance y datos a emplear. 8 – 10 puntos Se identifican parcialmente el problema, su contexto, pregunta de negocio, alcance y datos a emplear, o falta alguno por identificar. 6 – 7 puntos. No se identifican dos o más de los elementos solicitados. 0 – 5 puntos.
Modelos desarrollados y su evaluación.	<p>35 puntos</p> <ul style="list-style-type: none"> Se describe claramente el proceso de entrenamiento de los modelos, la selección de características y se evalúan adecuadamente los modelos con métricas apropiadas, todo debidamente soportado en MLflow. 28 - 35 puntos. La descripción del proceso de entrenamiento de modelos es parcial, su evaluación es limitada. 21 – 27 puntos. No es claro cómo se desarrolla el modelo, la selección de características o su evaluación. No hay soporte de MLflow. 0 – 20 puntos.
Observaciones y conclusiones sobre los modelos.	<p>10 puntos</p> <ul style="list-style-type: none"> Se destacan aspectos relevantes del comportamiento de los modelos y se concluye apropiadamente. 8 – 10 puntos. Se destacan algunos aspectos de los modelos, pero quedan faltando otros aspectos importantes. 6 – 7 puntos. No se destacan aspectos relevantes de los modelos o se hace de manera muy limitada. 0 – 5 puntos.
Tablero desarrollado.	<p>35 puntos</p> <ul style="list-style-type: none"> El tablero cumple con los objetivos planteados en la pregunta de negocio y alcance del proyecto, incluye las predicciones del modelo y su interfaz es adecuada para el usuario. 28 – 35 puntos. El tablero sigue los lineamientos de la pregunta de negocio y alcance del proyecto, incluye predicciones del modelo, pero su interfaz es limitada o poco adecuada para su uso. 21 – 27 puntos. El tablero no cumple con los lineamientos de la pregunta de negocio y alcance del proyecto, o no incluye predicciones del modelo. 0 – 20 puntos.

Entregable	Rúbrica
Reporte de trabajo en equipo.	<p>10 puntos</p> <ul style="list-style-type: none"> ■ El reporte de trabajo en equipo describe completamente las actividades de cada miembro del equipo y la distribución es adecuada y soportada. 8 - 10 puntos. ■ El reporte de trabajo en equipo describe completamente las actividades de cada miembro del equipo, pero la distribución no es adecuada. 6 – 7 puntos. ■ El reporte de trabajo en equipo describe parcialmente las actividades de los miembros del equipo. 0 – 5 puntos.