



**MIAD**

Maestría  
en Inteligencia  
Analítica de Datos



# Proyecto

## 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 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

Como el proyecto requiere un esfuerzo importante, se espera que las tareas asociadas al proyecto se vayan realizando a lo largo de las 8 semanas del curso. A continuación listamos las tareas principales que se espera se desarrollen cada semana. Estas actividades se consideran como las mínimas necesarias para garantizar un buen avance del proyecto.

### Semana 1

En esta primera semana debe por lo menos:

- Definir el equipo de trabajo para el proyecto.
- Explorar posibles problemas que el equipo pueda abordar en el proyecto.
- Para cada posible problema identificar claramente la pregunta de negocio y el conjunto de datos asociado, garantizando su disponibilidad.
- Documentar.

### Semana 2

En esta semana el equipo debe por lo menos:



- Definir el problema que abordarán en el proyecto. Para priorizar y seleccionar es importante considerar la disponibilidad inmediata de datos.
- Definir la pregunta de negocio, el alcance del proyecto y los conjuntos de datos a emplear.
- Identificar cómo puede resolver la pregunta de negocio a través de visualizaciones de los datos (descriptivo) y un modelo (predictivo).
- Elaborar una maqueta inicial del prototipo donde se identifiquen claramente sus elementos y su relación con la pregunta de negocio a resolver.
- Documentar.

### Semana 3

En esta semana el equipo debe por lo menos:

- Establecer repositorios en Git para el código y DVC para los datos.
- Realizar una exploración de los datos, versionando su código y datos.
- Iterar sobre la maqueta del prototipo de acuerdo con lo que vayan encontrando en los datos.
- Documentar.

### 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.



## Semana 6

En esta semana el equipo debe por lo menos:

- Desarrollar nuevas versiones de los modelos, comparar y seleccionar mejores alternativas. Emplear MLflow para versionar los modelos y los resultados de sus experimentos.
- Empaquetar y desplegar una primera versión del tablero, API y los modelos.
- Documentar.

## Semana 7

En esta semana el equipo debe por lo menos:

- Desarrollar nuevas versiones de los modelos, comparar y seleccionar mejores alternativas. Emplear MLflow para versionar los modelos y los resultados de sus experimentos.
- Empaquetar, integrar y desplegar una nueva versión del tablero, API y los modelos.
- Documentar.

## Semana 8

En esta semana el equipo debe por lo menos:

- Empaquetar, integrar y desplegar la última versión del tablero, API y los modelos.
- Documentar.
- Preparar entregables finales.

# Entregas

## Entrega 1

Al final de la semana 3 (**domingo 26 de octubre**) deberán entregar un reporte y soportes sobre los siguientes elementos:

- Problema que abordarán y su contexto.
- Pregunta de negocio y alcance del proyecto.
- Conjuntos de datos a emplear.
- Repositorio Git en uso para el código.
- Repositorio DVC en uso para los datos.



- Exploración de los datos.
- Maqueta (mockup) del prototipo.
- Reporte de trabajo en equipo.

## 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 (mockup).
- 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.

## Entrega 3

Durante la semana 8 (**martes 25 de noviembre**) deberán entregar un reporte y soportes sobre los siguientes elementos:

- Repositorio Git con todo el código desarrollado. Aquí deben reflejarse los aportes de todos los miembros a través de commits. Deben incluirse las versiones de los modelos, pipelines de entrenamiento y procesamiento de datos, así como las fuentes del tablero y los artefactos para empaquetar y desplegar el tablero.
- Datos versionados en DVC.
- Modelos desarrollados.
- Experimentos realizados y su soporte en MLflow.
- Modelos empaquetados y desplegados en la API.
- Artefactos para desplegar el tablero y el modelo.
- Manual de usuario del tablero.
- Manual de instalación del tablero.
- Reporte de trabajo en equipo.
- Un video en el cual se describa el trabajo realizado a lo largo del semestre. El video debe contener una síntesis del problema abordado, así como su relevancia, los modelos construidos, la solución (tablero) y, los principales resultados y conclusiones. El video no podrá exceder los 10 minutos de duración.



**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 y la **sustentación**.

## 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.