



Proyecto de IA y Machine Learning

Esta actividad le servirá para practicar técnicas de *machine learning*. Consistirá en desarrollar y documentar un proyecto mediante Python a partir de cualquier tipo de datos. En este caso, utilizaremos un conjunto de datos del ámbito empresarial.

El proyecto se realizará en dos fases

1. Diseño del proyecto: el equipo debe recopilar datos sobre una organización de su elección y analizar una implementación posible del modelo (los miembros del equipo deben tener conocimientos en la materia).
2. *Machine learning*: el equipo debe implementar el proyecto en Python y desarrollar las tareas de *machine learning* expuestas en el curso.

Instrucciones

- Se trata de un proyecto en equipo.
- Los equipos no deberían estar formados por más de **5** participantes.
- Todos los miembros deben contribuir a desarrollar el código base del proyecto. No olvide marcar de forma clara las secciones del cuaderno final en las que ha trabajado cada miembro.
- Si un participante no encuentra compañero de equipo, deberá ponerse en contacto con Academic Support.
- Todos los miembros del equipo deben ser responsables del progreso del proyecto.

Calendario

Módulo 1	Asignación de proyectos y formación de equipos	
Módulo 4	Entrega de las descripciones del proyecto	20 puntos
Módulo 8	Entrega del cuaderno de IPython, presentación del proyecto e informe final	80 puntos

Entregas

- Los informes deberán entregarse en el campus virtual.
- Cada equipo debe incluir un solo archivo en cada entrega. Uno de los miembros se encargará de enviarlos bajo el nombre de todo el equipo.

Calificación

- La puntuación total del proyecto son 100 puntos, una parte considerable de la calificación final.

Evaluación

Cada participante será evaluado de forma individual en función de su contribución al código base del proyecto. Los criterios de evaluación son los siguientes:

- Claridad: ¿la estrategia de datos de la organización se explica con claridad?
- Originalidad: ¿la organización elegida es lo suficientemente innovadora?
- Contenido: ¿el modelo tiene suficiente contenido (datos no estructurados, algoritmos de *machine learning*, visualización, etc.)?
- Corrección: ¿el diseño del proyecto es sólido y está bien elegido?

Ejemplos de proyectos

Kaggle.com es un buen punto de partida. Además, se pueden tomar como referencia algunos de los mejores proyectos que hasta ahora se han presentado en los campos de *machine learning*, *data science* y entrenamiento de *deep learning*:

- Análisis de grafos de Divvy Bikes
- Predicción de precios de Airbnb
- Predicción nutricional de Wholefoods
- Predicción de salario mediante Indeed y Glassdoor
- Predicción de solicitudes de visados H-1
- Clasificación de café de Starbucks
- Detección de marcas de moda
- Pronóstico de decisiones de inversión
- Análisis de la transmisión de la COVID-19

También se pueden emplear solo las técnicas que hemos visto en el curso para garantizar una implementación sólida.

Entrega en el módulo 4: lista de miembros del equipo, tareas de cada miembro y descripción del caso (mínimo 1 página con interlineado sencillo).

Entrega en el módulo 8: informe final (mínimo 10 páginas con interlineado sencillo, sin contar la bibliografía), documento de Jupyter Notebook con un archivo de nombre "predicciones.csv" (si corresponde) y presentación de PowerPoint de 10 diapositivas como máximo.

Durante la presentación del proyecto, los miembros del equipo deberán estar preparados para responder a cualquier pregunta sobre su proyecto. Asimismo, tendrán que vestirse con arreglo a los códigos estéticos de un entorno profesional.