

Machine learning con Python

M^a José Gómez Silva
Luis Garmendia Salvador
José Ángel Carballo Sánchez

ÍNDICE

Los contenidos de la asignatura serán los siguientes:

SEMANA 1:

1. **Introducción:** durante este módulo veremos conceptos básicos que después necesitaremos conocer durante las clases. También aprenderemos a utilizar la herramienta jupyter notebook y realizaremos un repaso general del uso en las principales librerías de data science de Python.

- Introducción al machine learning y sus aplicaciones.
- Conceptos clave: modelo, fases de entrenamiento, validación y test, función de coste.
- Tipos de aprendizaje.
- Problemas del aprendizaje automático.
- Bases para implementar un proyecto de ML: python, jupyter, scikit-learn, pandas.

2. **Ingesta de datos:** en este apartado, veremos casos reales de obtención de datos y como convertirlos a dataframes.

- Tipos y fuentes de datos: tipos de archivos, dimensiones, estructurados y no estructurados, anotaciones.
- Organización y tabulación de los datos: dataframe, manejo de filas y columnas.
- Visualización y análisis de los datos (matplotlib).

3. **Pre procesado:** veremos diferentes técnicas de procesado y limpieza de datos.

- Filtrado
- Balanceo
- Aumento de datos

SEMANA 2:

4. **Modelos:** aprenderemos mediante ejemplos los algoritmos supervisados y no supervisados más conocidos: Regresión - Clasificación - Clustering

Entre los algoritmos supervisados haremos regresión y los clasificadores más conocidos como:

- Naive Bayes. Clasificación de textos.
- SVM
- Árboles de decisión
- Random Forest. Clasificando dígitos.
- Red Neuronal de una capa (regresión logística)

Entre los algoritmos no supervisados de clustering veremos los fundamentos y haremos ejemplos de:

- K-Means
- PCA

SEMANAS 3 Y 4:

5. **Mejora de modelos:** una vez aprendido las técnicas de importado, procesado, limpieza de datos y los algoritmos que debemos utilizar (según el caso de uso que necesitemos implementar), entramos dentro del apartado de mejora de modelos. Entre las diferentes técnicas, utilizaremos: feature engineering, selección de variables...
6. **Automatización de modelos:** muchas veces, necesitaremos probar entre los diferentes algoritmos cual es la combinación óptima que mejor funciona. Aprenderemos a utilizar técnicas , herramientas y librerías que nos ayudarán a ello.
7. **Despliegue de modelos:** tras haber aprendido a utilizar todas las tecnologías durante las clases, veremos las diferentes maneras de entregar un modelo y ponerlo en producción.
8. **Ejercicios:** por último, terminaremos la asignatura con la realización de diferentes ejercicios y la práctica de evaluación.

INSTALACIÓN

La clase de machine learning con Python la seguiremos a través de la herramienta "Jupyter" a través de "Notebooks", que permiten una fácil visualización y flexibilidad a la hora de hacer exploración de datos y modelos de manera ágil.

Realizaremos la instalación (antes de la clase) de esta herramienta (no es necesario instalar jupyter lab, solamente jupyter notebook):

<https://jupyter.org/install>

Si ya lo tenéis instalado, no es necesario nada extra (quizás hagamos algún ajuste de versionado de librerías durante las clases)

EVALUACIÓN

La evaluación de la asignatura Machine learning con Python será de la siguiente manera:

- Realizar el caso práctico que se presentará junto al resto de materiales (4 horas de realización)