

PROYECTO

El proyecto consistirá en la realización de una segunda versión del código desarrollado en la primera parte de la materia (Aprendizaje Automático I), para el mismo conjunto de datos, pero ampliando el proyecto en los siguientes aspectos:

1. Uso de al menos una técnica, a elegir, entre las siguientes:
 - a. Técnicas de reducción de la dimensionalidad de los datos (PCA, ICA, etc.)
 - b. Técnicas de agrupamiento (clustering) para plantear el problema inicial como un problema de agrupamiento, sin considerar, por tanto, las etiquetas de clase para el entrenamiento de este modelo.
 - c. Técnicas de detección de anomalías para abordar el problema inicial como un caso de detección de anomalías. Considerando una, o varias, de clases como la clase normal y la otra/s como anomalías.
2. Ampliación del código para que permita realizar mediciones de tiempos de entrenamiento medios de los modelos de aprendizaje automático desarrollados para la clasificación supervisada.
3. Uso de programación multiproceso o multihilo para acelerar la ejecución de alguna parte del código, por ejemplo, la ejecución paralela en varios cores de los k experimentos de la validación cruzada. Se podría comprobar el tiempo de CPU necesario para ejecutar esa parte del código de forma secuencial o en paralelo.

Para el informe a entregar, se podrá usar como base el entregado para el proyecto anterior añadiendo simplemente aquellos aspectos incluidos en esta segunda versión del código:

- Resultados obtenidos con la reducción de la dimensionalidad, el clustering o la detección de anomalías.
- Tiempos de ejecución de los modelos entrenados con un único proceso o con multiproceso.

Documentación a entregar

- Fichero/s con el código fuente del proyecto en Python. El código deberá incluir comentarios para documentar las partes fundamentales del mismo.
- Informe del proyecto, preferiblemente en formato Word o pdf, que incluya la descripción las ampliaciones realizadas y resultados obtenidos.

Datos y fechas importantes

- El proyecto se desarrollará con la misma pareja del proyecto de Aprendizaje Automático I o de forma individual.
- Esta actividad supone un 50% de la nota final de la asignatura.
- Fecha y hora límite de entrega: **18 de mayo de 2021 a las 23:59h.**
- Lugar de entrega: enlace de la tarea en el [Campus Virtual](#) de la UDC.

Criterios de evaluación del proyecto:

- Calidad del código fuente entregado. El código deberá incluir todos los aspectos mencionados en los apartados anteriores.
- Calidad de la memoria del trabajo. Se valorará especialmente la amplitud del estudio experimental realizado, el análisis de los resultados y las conclusiones del trabajo. También se valorará la claridad del documento y de las explicaciones.
- Si se detecta algún plagio entre los códigos fuente entregados, la nota de esta actividad será de un 0 para todos los grupos implicados. Para la detección de plagios se empleará una herramienta automática.

En Ferrol, a 30 de abril de 2021.