Hackaton

“Tackling a real-world challenge in Industry 4.0: NASA Turbofan Jet Engine Dataset”

“Enfrentándonos a un reto del mundo real en la Industria 4.0: conjunto de datos de motores *turbofan* de la NASA”

Organizado como actividad del programa de doctorado “Computación avanzada, energía y plasmas” (Universidad de Córdoba)

Prognostics and health management is an important topic in industry for predicting state of assets to avoid downtime and failures. We will face a Kaggle version[1] of the very well known public data set for asset degradation modeling from NASA. It includes Run-to-Failure simulated data from turbo fan jet engines. The goal is to predict the remaining useful life (RUL) of each engine in the test dataset, which is the number of remaining flights. We will have the opportunity to answer the following questions:

- *Can we efficiently predict the remaining useful time for the engine (error significance)?*
- *Can we improve state of the art results?*
- *Which features are the most important for predicting the failure of the turbofan engine?*
- *Does adding historical data improve our model?*
- *Is the collected data sufficient enough to give an accurate prediction?*
- *How can we turn our problem to a classification one?*

We will work with Python, Anaconda, Google Colab, Kaggle, numpy, pandas, scikitlearn, tensorflow, seaborn, matplotlib...

El pronóstico y la gestión son temas importantes en la industria para predecir el estado de los activos y evitar tiempos de inactividad y fallos. Nos enfrentaremos a una versión Kaggle[1] de un conjunto de datos públicos muy conocido para el modelado de la degradación de activos de la NASA. El objetivo es predecir la vida útil restante (RUL) de cada motor en el conjunto de datos de prueba, que equivale al número de vuelos que le quedan al motor. Tendremos la oportunidad de responder a las siguientes preguntas:

- ¿Podemos predecir eficazmente el tiempo útil restante del motor (importancia del error)?
- ¿Podemos mejorar los resultados del estado del arte?
- ¿Qué características son las más importantes para predecir el fallo del motor *turbofan*?
- ¿La adición de datos históricos mejora nuestro modelo?
- ¿Son suficientes los datos recopilados para ofrecer una predicción precisa?
- ¿Cómo podemos convertir nuestro problema en un problema de clasificación?

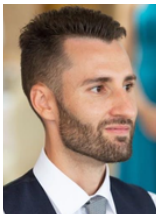
Trabajaremos con P*ython, Anaconda, Google Colab, Kaggle, numpy, pandas, scikitlearn, tensorflow, seaborn, matplotlib…*

Jueves 30 de marzo de 2023, de 10.30h a 12.30h, Sala S1 del edificio Ramón y Cajal, Campus de Rabanales.

Ponente: **Luca Romeo** (Doctor en Informática)

University of Macerata (UniMC, Macerata, Italia)

https://docenti.unimc.it/luca.romeo



LUCA ROMEO received a Ph.D. degree in computer science from the Department of Information Engineering (DII), Università Politecnica delle Marche, in 2018. His Ph.D. thesis was on “applied machine learning for human motion analysis and affective computing”. He is currently a Tenure Track Assistant Professor of Computer Science with University of Macerata | UniMC Department Economics and Law. He is also Adjunct Professor of Customer Intelligence & Big Data, at Luiss, Roma and he is affiliated with the Unit of Computational Statistics and Machine Learning, Fondazione Istituto Italiano di Tecnologia Genova. His research topics include the design of novel Machine learning algorithms for solving relevant challenges in different real-world domains.

1 NASA Turbofan Jet Engine Dataset https://www.kaggle.com/datasets/behrad3d/nasa-cmaps