Pablo Ariño Fernández

Madrid, Spain pablo.arino.25@ucl.ac.uk pabloo22.github.io pablo-arino pabloo22

"In times of change, learners inherit the earth, while the learned find themselves beautifully equipped to deal with a world that no longer exists." — Eric Hoffer

Education

M.Sc. **Machine Learning**, Univeristy College London — Sept 2025 – Sept 2026

- Fully funded by the"La Caixa" Foundation. Awarded to 100 students across Spain and Portugal to pursue international studies in any discipline.

B.Sc. **Data Science and Artificial Intelligence**, Polytechnic University of Madrid — Sept 2020 – June 2025

- **GPA:** 93% (best academic record in Data Science and Artificial Intelligence award & 13 honour grades)
- **Thesis**: Solving the Job Shop Scheduling Problem with Graph Neural Networks: A Customizable Reinforcement Learning Environment (10/10, 2nd prize in CECVImpulsaIA360)
- Tutor for incoming students during the first semesters of 2021 and 2022

Exch. **ERASMUS Exchange Program**, Poznań University of Technology — Feb 2023 – Jun 2023

- **GPA:** 97.4%
- **Coursework:** Computational Intelligence, Robotics, Advanced Natural Language Processing, Decision Analysis, Data Visualisation, Polish

Exch. **ATHENS Exchange Program**, Ecole des Ponts ParisTech — Nov 2022

- **Coursework:** Uncertainty Quantification in Numerical Simulations

Experience

Spanish National Research Council (CSIC), JAE Intro Research Fellow — Oct 2024 – Apr 2025

- Identified challenges in explainability techniques applied to the state-of-the-art deep learning models for phylogenetic trees
- Studied graph-based encodings for phylogenetic trees to predict key parameters and presented my results at the MCEB Congress.
- Participated in the XVII scientific marathon as a co-author of the presentation "AI and its capacity to understand the generation and extinction of species"
- Supervisor: Isabel Sanmartin

AI+DA Research Group, Research Fellow — Sep 2023 – Oct 2023

- Conducted a comprehensive literature review, focusing on the application of AI and Large Language Models to enhance learning and educational methods
- Supervisor: Alejandro Martín
- Learnt about the science of learning through research articles and the books *Make it Stick* and *Ultralearning*

MÁSMÓVIL Chair, Intern Data Scientist — April 2022 – Jan 2023

- Utilised advanced models (Transformers, XGBoost, and ensemble methods) for churn prediction using an imbalanced dataset
- Applied Bayesian optimisation and genetic algorithms for hyperparameter optimization using a custom metric to adapt model selection to the company's needs
- Created data analysis reports for data stored in multiple tables and more than one thousand columns in total. Some of them included visualizations with t-SNE

Achievements

First Ascent Spain participant. Classified among more than one thousand candidates to participate in the trip to Milan organized by Bending Spoons. LinkedIn post

Second place in WhiteBox's Datathon. Achieved 2nd place in a Kaggle-style competition focused on vehicle valuation modelling, leveraging xgboost on real-world data provided by DataMarket, with less than 3 hours to clean data, engineer features, and train the model. LinkedIn post

Winner of Google's Developer Student Club Datathon. Questionnaire on a dataset with 30 questions. First place with maximum score. LinkedIn post

Sample-based Learning Methods Certificate. Issued in Oct 2021 by the University of Alberta. Part of the Reinforcement Learning specialization. Audited the other courses. Certificate

Competitive Programming. Selected to represent the Polytechnic University of Madrid in the Ada Byron contests of 2021 and 2022. Teams of 2021 Teams of 2022

Winner of Kaggle-like competition about extracting characters from CAPTCHA images. It was part of the Machine Learning II course and had two phases: a preparatory phase, and an in-person phase. github link

- Used a convolutional vision transformer (ResNet + transformer encoder)
- Used Weights&Biases for experiment tracking
- Followed the project template suggested in the Cookiecutter Data Science template

Projects

Library for the Job Shop Scheduling Problem (JSSP) — Feb 2024 - Present — github

- Includes a highly customisable reinforcement learning gymnasium environment to solve the problem sequentially with graph neural networks
- Followed good software engineering practices: modular design (SOLID and *Clean Code*'s principles), automatic tests, linting, and documentation page
- Being used by the community: 58 stars and eight contributors.

Generating Realistic and Difficult JSSP Instances with an Adapted Adjacency-Matrix-based Generative Adversarial Neural Network — Nov 2023 – Jan 2024 — github

- Wrote a research article draft for the Data Science Project course
- Proposed a theoretical framework that increased the size of the generated graphs by an order of magnitude while using the same computational resources
- Introduced a novel metric for evaluating the difficulty of solving an instance

Retrieval-based Question Answering System for videos — May 2023 - June 2023 — github

- Implemented both retrieval augmented generation and extractive question answering methods for answering questions about YouTube playlists
- Developed with LangChain, Hugging Face, and Streamlit

Vectorised Tabular Reinforcement Learning Framework — Dec 2022 - Jan 2023 — github

- Includes dynamic programming and double Q-learning methods
- Solved the Jack's Car Rental problem (see a demo)

Keras-like Deep Learning Framework from Scratch (with NumPy) — Nov 2021 - Jan 2022 — github

- Created the Sequential model with Dense, Conv2D, Dropout, and Flatten layers
- Supports multiple optimisers (SGD and SGDMomentum), multiple activation functions, including softmax; and classic cost functions

Technologies and Tools

Languages: Python, SQL, Matlab, R, C, HTML, LaTeX. **Software development:** Docker, Git. **Deep Learning:** PyTorch, TensorFlow, Keras, PyTorch Geometric. **ML libraries:** scikit-learn, sktime, xgboost. **Optimisation:** ORTools, Optuna.