

Alberto Agudo Dominguez

✉ a.agudo99@hotmail.es in LinkedIn 📄 github.com @ stackoverflow 📝 blog/personal website

Current role	Data Scientist at Ebiquity
Experience	Analytics, Media, Financial Services, Governments, Artificial Intelligence, Consulting, Energy, Sustainability
Technologies	AWS, Python, R, Spark, Git, SQL, Power BI

Right now I'm learning...
about the foundations and applications of LLMs as well as MLOps for their deployment, particularly in AWS as I received the AWS Solutions Architect – Associate badge in July 2023

Work experience

Data Scientist, Ebiquity 🏢 Sep 2022 – Present (1y 5m)

Analytics · Media

AWS Python R Spark Git

- Reduced project turnaround time by 33% through the development of new statistical tools that enhanced consultants' workflows
- Decreased data review and presentation time by two working days per project by creating Shiny apps for consultants to quickly explore and visualize their data
- Accelerated data processing and improved script efficiency, resulting in 2x to 100x faster data preparation through Spark optimizations in AWS and Arrow
- Conducted end-to-end projects valued at over GBP 100k through data collection and model development. These models included xgboost for customer choice classification and brand performance prediction, as well as econometric models for causal inference
- Successfully built and delivered an ML POC which resulted in new revenue streams through Xgboost

Visiting Researcher, Bank of Spain 🏢 Jun 2022 – Sep 2022 (3m)

Financial Services · Governments

Python R

- Created a comprehensive scholars' database to support the development of a research paper
- Developed and implemented a web scraping tool utilizing Selenium to collect academic data from Google Scholar
- Utilized the Twitter API to extract relevant tweets from top economics departments in the United States

Data Science Internship, Smartup Digital Agency 🏢 Feb 2021 – Jul 2021 (5m)

Artificial Intelligence · Consulting

SQL Python Power BI

- Delivered three end-to-end machine learning projects worth more than EUR 50k each, focused on geospatial crime analysis, demand forecasting and marketing effectiveness analysis, and machine optimization

- Utilized Python (Random Forests, Prophet) and PowerBI to perform statistical modeling, machine learning, programming, and dashboard creation to solve complex business problems, presenting results

Research assistant, University of Salamanca

Nov 2019 – Feb 2021 (1y 3m)

Energy · Sustainability

Python

R

- Contributed to the ATACC research project, funded by the Ministry for Ecological Transition, focused on adaptive policies for water management under climate change
- Published a paper in a Q1 peer-reviewed academic journal on water management as corresponding author
- Created databases and conducted econometric modeling in R

Education

The London School of Economics and Political Science (LSE)

Sep 2021 – Sep 2022 (1y)

(Distinction) MSc Applied Social Data Science, Data Science

- Expertise in Computer Programming, Data Engineering, Causality, ML, and Deep Learning. Specialized in Deep Learning, foundations of GenAI and LLMs (Transformers, GANs)
- Technologies used: Python, R, AWS, SQL, Tensorflow
- Awarded with one of the ten "Postgraduate Studies" grants of the Bank of Spain (former Spain Central Bank) to attend the LSE MSc

University of Salamanca

Sep 2017 – Jul 2021 (3y 10m)

Bachelor (9.73 / 10) Economics

- Top of the class award (28 subjects with honours)
- Erasmus Exchange at KU Leuven (Belgium), completed with distinction
- Honours in the Thesis Defense

Snippets

Deep learning project on NLP experimentation @ [github.com](#)

Performed sentiment analysis of movie reviews. Compared 80+ deep learning architectures for best results based on a combination of different types of embeddings through vector databases (training our embeddings vs other popular options like SpaCy) and standard architectures, like LSTMs/GRUs/RNNs. Developed a suite of tools for text preprocessing (tokenizers) and semantic representation.

Project on Python software for tennis data analysis @ [github.com](#)

Created several modules to process tennis data and analyze it. The main objective was to create different ranking algorithms and compare their performance with the WTA rankings, without relying on relevant dependencies like NumPy / Pandas to prove code efficiency.

Publication as a first author at a Q1 journal on water management @ [www.sciencedirect.com](#)

Publication as a first author as a result of my work as a research assistant at the ATACC project (University of Salamanca), where I created econometric models to predict the impact of climate change on future droughts, and its consequences for the feasibility of a drought insurance scheme for irrigation farmers.