Juan Arturo Abaurrea Calafell

juanarturoabaurreacalafell@gmail.com | github.com/Jarturog | linkedin.com/in/j-arturo-a-c

Education

Polytechnic University of Madrid *Sep 2025 – Jun 2026, Madrid, Spain*
Master's Degree in Machine Learning and Big Data. Class representative.
- **Status:** Currently enrolled.
- **Relevant Coursework:** Research Methodology and Techniques, Evolutionary and Bio-inspired Computing, Deep Learning for Natural Language Processing, Deep Learning for Computer Vision, Reinforcement Learning and Generative Techniques, AI Models for Time Series.

Beijing Institute of Technology *Jan 2026 – Jan 2026, Zhuhai, China*
BIT (Zhuhai) Winter Program - Artificial Intelligence.
- **Relevant Coursework:** Unsupervised and Supervised Learning, Convolutional Neural Networks, Recurrent Neural Networks, Graph Neural Networks, Reinforcement Learning.

University of the Balearic Islands *Sep 2021 – Jun 2025, Palma de Mallorca, Spain*
Bachelor's Degree in Computer Engineering – Specialized in Artificial Intelligence and Computing.
- **Grade:** 8.3 out of 10 (8 honors).
- **Relevant Coursework:** Machine Learning (9.2), Artificial Intelligence (honor), Intelligent Systems (9.0), Data Mining (7.0), Advanced Algorithms (honor), Databases I & II (7.0 and honor), Statistics (8.5).

Ulm University of Applied Sciences *Mar 2024 – Jul 2024, Ulm, Germany*
Exchange semester. Completed Bachelor's thesis and three additional subjects.
- **Grade:** 9.7 out of 10.
- **Bachelor's Thesis (9.6):** Implemented a chatbot for colonoscopy preparation using NLP techniques.

Experience

TravelgateX, Entry Software Engineer *Oct 2024 – Dec 2024, Palma de Mallorca, Spain*
Backend development in C# and Visual Basic using Visual Studio, GitHub, Fiddler, Postman, Elasticsearch, and Grafana.

Languages

Native: Spanish, Catalan | **C1:** English (Cambridge certified) | **A1:** German

AI & Data-Related Skills

- **Programming Languages:** Python, R, SQL (MySQL, MariaDB, PostgreSQL), MATLAB
- **Machine Learning & AI:** PyTorch, Scikit-learn
- **Data Science:** NumPy, Pandas
- **Data Visualization:** Matplotlib, Seaborn, Plotly, Dash, Grafana
- **Tools & Platforms:** Docker, Git, GitHub Actions, Azure, Jupyter Notebook

Relevant Projects

Cybersecurity AI Datathon - 2nd Place *Nov 2025*
- Competed in an AI datathon focused on applying machine learning to cybersecurity and ransomware attack classification.

- Achieved an aggregate F1-score of 0.88424 across two classification problems, narrowly behind first place (0.88490).
- Applied prompt engineering and jailbreaking techniques to extract three passwords from a custom LLM that were required to decode encrypted dataset columns.

Image Classification, Detection and Segmentation Using Neural Networks *Dec 2024 – Jan 2025*
- Implemented and compared five deep learning models (ResNet50, VGG16, Inception v3, and two custom CNNs) for image classification, achieving 90-96% F1-score with a custom architecture on a dataset of less than 100 samples per class.
- Designed a lightweight custom CNN architecture (625K parameters) achieving 8-20× faster inference than pretrained models while maintaining comparable performance with minimal training data.
- Developed object detection and segmentation pipelines using YOLOv8m and Mask R-CNN, performing fine-tuning experiments and comparative analysis of pretrained versus adapted models.
- Applied explainability techniques (SmoothGradCAM++) to interpret model predictions.

Real Estate Price Analysis and Prediction Using Data Mining *Dec 2024 – Jan 2025*
- Preprocessed large-scale dataset (114K samples, 25 features) using IQR outlier detection, Random Forest imputation, and feature engineering techniques.
- Developed predictive models (K-NN, Random Forest, Linear Regression) for property construction year, achieving <5% error with optimized K-NN.
- Performed clustering analysis (K-medoids, hierarchical methods) to investigate relationships between construction year, energy efficiency, and conservation status.
- Quantified coastal location impact on property prices through statistical analysis (Welch's t-test) and regression modeling, isolating a €97K price increase attributable solely to coastal proximity.

Chatbot for Colonoscopy Preparation *Mar 2024 – Nov 2024*
- Conducted research on chatbots for colonoscopy preparation, identifying key limitations in existing approaches.
- Developed a novel solution incorporating multiple NLP techniques including Levenshtein distance, Jaccard index, TF-IDF, BM25, and sentence embeddings to address identified gaps in the field.

Other Technical Skills

- **Programming Languages:** Ada, Assembly, Bash, C, C#, C++, CSS, GoLang, HTML, Java, JavaScript, Lisp, PHP, Prolog, Visual Basic
- **Tools & Technologies:** Bootstrap, Docker, Fiddler, Git, GitHub, HubSpot, Insomnia, Jira, NetBeans, Postman, PyCharm, Visual Studio, Visual Studio Code, Wireshark, XAMPP

Soft Skills

- **Initiative:** Demonstrated through my role as class representative and my exchange in Germany.
- **Rational Thinking:** Methodical problem-solving approach supported by my academic performance.
- **Curiosity and Continuous Learning:** Constantly exploring new ideas and technologies, as reflected in my academic path.
- **Perseverance and Patience:** Successfully completed my first large-scale solo project (my bachelor's thesis) with top marks, and subsequently balanced a full-time internship in the mornings with university classes in the afternoons, demonstrating strong commitment and resilience.
- **Teamwork and Communication:** Strengthened collaboration and communication through numerous group projects, such as with my image processing project and data mining assignment, as well as leadership experience as class representative.