

ANA-MARIA LUISA MOCANU

Data Scientist

@ m.mogoase23@gmail.com ☎ +40 738 434 378
📍 Bucharest - Romania 🔗 <https://github.com/medz23>
in <https://www.linkedin.com/in/ana-maria-luisa-mocanu-87b4221b3/>



EXPERIENCE

Data Scientist

Endava

📅 Jun 2026 – present 📍 Bucharest, Romania

- Applying machine learning and NLP to data-driven problems, building on a strong software-engineering background from previous fintech projects.

Software Engineer

Endava

📅 Dec 2021 – Jun 2026 📍 Bucharest, Romania

- Enhanced my skills in software development, Agile methodologies, and attention to detail while working with service-oriented and microservice architectures.
- Built and maintained fintech applications from scratch in Java and Spring Boot across multiple projects, one nominated as one of Endava's most successful projects.
- Grew from independent feature work into owning automation testing and DevOps tasks, deployments, gateway (Kong) configuration, and release preparation, in close collaboration with clients and product owners.
- Gained experience with the following technologies: Java, Spring Boot, JavaScript, Angular, HTML, CSS, Cosmos, Redis, MongoDB, Mongock, Liquibase, PostgreSQL, Docker, Kafka, AWS, Azure, GitHub, Bitbucket, Maven, Jira, Bruno, Linux.

Python Teaching Assistant

University Politehnica Bucharest

📅 Oct 2025 – present

- Teach beginner-level Python to Bachelor's students in Google Colaboratory, guiding them through lab exercises, projects, and tests, with positive feedback for making programming concepts easy to understand.

PROJECTS

IROH: Humor Retrieval Pipeline (JOKER 2026, CLEF)

- Built a three-stage humor retrieval pipeline that combines hybrid sparse-dense retrieval, a finetuned cross-encoder reranker, and an ensemble of LoRA/QLoRA-adapted LLM judges, ranking 1st on the official JOKER 2026 Task 1 English leaderboard with a MAP of 0.6347.
- Trained the humor judges by distilling query-aware rationales from a vision-language model, and generated structured hard negatives to help the reranker tell real jokes apart from look-alikes. Link: github.com/DS4AI-UPB/VANGUARD-CLEF2026-JOKER

EXIST: Multimodal Sexism Detection in Memes (CLEF 2026)

- Developed a human-centered multimodal model that fuses five inputs, the meme text, a generated visual description, the image, the annotator demographics, and their physiological signals, grounding the description in the image with cross-attention and conditioning on the annotator context with FiLM.

EDUCATION

PhD in Computer Science

University Politehnica Bucharest

📅 Oct 2025 – present

Master in Artificial Intelligence

University Politehnica Bucharest

📅 Oct 2021 – Sep 2025

Bachelor in Mechanical Engineering

University Politehnica Bucharest

📅 Oct 2017 – Jul 2021

TECHNICAL SKILLS

- **Programming & Development:** Java, Python, Bash, Spring Boot, Spring MVC, Django, Flask, Android, Angular, HTML, CSS, JavaScript, Maven, Gradle, Tomcat.
- **Machine Learning & Data:** AI, ML, Deep Learning, NLP, Data Science, PyTorch, TensorFlow, Keras, Transformers, Hugging Face, spaCy, OpenCV, Pandas, NumPy, Matplotlib.
- **Databases:** MySQL, PostgreSQL, SQLite, Azure Cosmos DB, Redis, MongoDB.
- **Tools:** Docker, Microservices, Kafka, CI/CD, Terraform, Selenium, Swagger UI, Postman, Liquibase, Mongock.
- **Cloud & Systems:** Azure, AWS, Linux, WSL, SSH, Kubernetes (basic).
- **Collaboration:** Git, GitHub, Bitbucket, Jira, Agile.

PERSONAL SKILLS

- Critical Thinker, Adaptability, Good Interpersonal Skills, Time Management, Team Player, Problem Solver.

HOBBIES

- Learning new things, Sudoku, Listening to Music, Hiking.

- Framed the task as label distribution learning with a soft-label KL divergence loss and a supervised contrastive auxiliary objective, ranking 29th of 114 on Subtask 2.2 under soft evaluation. Link: github.com/DS4AI-UPB/VANGUARD-CLEF2026-EXIST

Aspect-Based Sentiment Analysis for Romanian Product Reviews

- Created the Romanian Review (RoR) dataset with 30,000+ e-commerce product reviews, featuring manual (3,000+) and automated annotation for ABSA across 14 product aspects.
- Built a hybrid annotation framework combining manual labeling with LLMs (GPT-3.5, GPT-4, Claude Haiku) for automatic labeling, producing the first Romanian ABSA dataset with mixed aspects and multiple sentiments.
- Distilled LLM annotations into lightweight (~100M-parameter) Romanian transformers (RoBERT and BERT Romanian cased), with the GPT-4.1-supervised students reaching F1 0.787–0.788, on par with their teacher and well ahead of GPT-3.5 Turbo and Claude Haiku variants.

ACAT: Collaborative Annotation Tool for ABSA

- Built a web-based annotation platform (Flask, PostgreSQL, Docker) covering four ABSA tasks of increasing granularity, computing inter-annotator agreement directly into the export step and exporting to CSV, JSON, and XML. Paper: [arXiv:2606.04189](https://arxiv.org/abs/2606.04189)

Python Exam IDE - Automated Grading Platform

- Built a web-based exam platform where students solve Python problems in a browser IDE (Monaco editor) and their code is graded automatically inside sandboxed Docker containers with memory and CPU limits, supporting both function-based and stdin/stdout grading modes.
- Implemented a professor dashboard (problem and group management, test cases, randomized problem pools, CSV export) and anti-cheat measures (fullscreen enforcement, tab-switch strikes, disabled copy/paste). Technologies used: Python, Flask, PostgreSQL, Docker, JavaScript. Link: github.com/medz23/scholar-rank

Abusive Language Detection

- Trained CNN and transformer models with various embeddings for hate-speech detection, achieving results that exceeded state-of-the-art performance on some datasets. Implemented ensemble methods using soft and hard voting and validated on established hate-speech benchmarks. Link: github.com/brittleru/abusive-language-detection

ACHIEVEMENTS

- Claude Code in Action – Anthropic – Jul 2026
- Building with the Claude API – Anthropic – Jun 2026
- Oral presentation – The 28th International Conference on Big Data Analytics and Knowledge Discovery – Aug 2026
- Oral presentation and Poster – The Conference and Labs of the Evaluation Forum – Sep 2026
- Poster Session – Best Romanian AI Thesis Awards – Romanian AI Days – Sep 2026
- Poster Session – Romanian AI Days – Sep 2025
- Applied Machine Learning – Google Digital Stack – Jun 2023
- Participated in NitroNLP hackathon – 2023
- NLP with Probabilistic Models – DeepLearning.AI – Jan 2022
- NLP with Classification and Vector Spaces – DeepLearning.AI – Jan 2022
- Android Fundamentals – Google Digital Stack – Jul 2021
- Python – Google Digital Stack – Feb 2021
- Volunteer for environmental activities: 'Caravana Voluntarilor' (2014–2015), 'TEAM' (2014–2016), 'Act for Tomorrow' (2019)