Maria **Salop**

Profile

Third-year Applied Data Science & AI student with expertise in transforming complex data into actionable insights using Python, SQL, and Power BI. Proficient in developing natural language processing solutions, including Transformers and speech-to-text technologies. Demonstrated ability to leverage data analytics for informed decision-making and enhance operational efficiency.

Skills

- Python and SQL (pandas, NumPy, scikit-learn)
- Data visualization (Power BI, matplotlib, plotly)
- Statistical analysis (descriptive stats, correlations)
- Data workflows (cleaning, feature engineering, ETL)
- Machine learning (classification, model evaluation)
- Natural language processing (Transformers, text preprocessing)
- MLOps and demos (FastAPI, Docker, Streamlit)
- Collaboration tools (Git/GitHub, Jupyter)
- Web technologies (HTML)

Contact details

@ salopmaria37@gmail.com
+31 621 613 890
linkedin.com/maria-salop
github.com/MariaSalop
Breda, North Brabant, Netherlands

Personal information

Citizenship: **Romania**

Languages: **English** (C1), **Romanian** (native)

Education

Applied Data Science & Artificial Intelligence: Academy of AGM. **2023–present**
Breda University of Applied Sciences.
◇ Machine Learning, Deep Learning, Data Science, AI development

Projects

AI & Data Science Project Portfolio *(Academic)* **09.2023 – 06.2025**
◇ A curated portfolio of academic and applied projects completed as part of my university education in Data Science and AI.
Link to the portfolio Link to the project

Air Pollution Mortality *Dashboard*
◇ Designed an interactive Power BI dashboard on global deaths from air pollution; performed data cleaning, correlation analysis, and data storytelling to surface key insights.
Link to the project

Player Valuation Model *NAC Breda*
◇ Built a regression model to estimate football player's market value from performance features; delivered an APA-style ethics report and model evaluation.
Link to the project

Waste Classification *CNN Project*
◇ Trained a CNN to classify recyclable waste; shipped a small demo app and ran user testing to validate usability and error cases.
Link to the project

Transport Incident Prediction *ANWB*
◇ Predicted traffic incidents using weather-linked features; accessed/joined datasets with SQL, engineered features, and compared baseline models.
`Deliverables not publicly available (NDA)`

Root Length Prediction *NPEC*
◇ Developed an image processing pipeline using OpenCV to measure root length from plant images. Included pre-processing, segmentation, and analysis.
Link to the project

Emotion Classification *NLP Research Project*
◇ Collaborated on an NLP research project to classify emotions in YouTube dialogue. Focused on improving accuracy and F1-score through class balancing, custom annotation, and error analysis. Deployed and evaluated the final model on real-world test data.
Link to the project

Chatbot Empathy Study *Digiwerkplaats*
◇ Co-authored a policy paper on chatbot empathy; wrote an individual research paper measuring effects on trust/comfort from a Qualtrics survey.
Link to the project

Emotion Classification *Deployment Project*
◇ Built a production-ready NLP pipeline (STT → preprocessing → optional translation → RoBERTa classifier). Exposed via FastAPI, packaged with Docker, deployed on Azure ML, and demoed in Streamlit.
Link to the project

Predictive Truck Maintenance System *In Progress*
◇ Building an end-to-end ML pipeline to predict truck component failures using 20 years of maintenance data. Integrating multiple data sources (SQL, PDF diagnostics, Excel, RDW) with ensemble models and SMOTE, targeting 10-15% cost reduction.