

# SAM STREUMER

Utrecht, Netherlands | +31 06 51310925 | Sam.streumer@gmail.com | [LinkedIn](#) | [samstreumer.com](#)

## EDUCATION

---

**Saxion University of Applied Sciences**, Apeldoorn, Netherlands

Expected 2028

- *B.S. in Applied Data Science & AI Engineering*
- *Study Program Committee Member*: Providing student feedback to shape curriculum for a newly launched program

## TECHNICAL SKILLS

---

- *Languages*: Python, SQL
- *Frameworks & Libraries*: scikit-learn, TensorFlow, PyTorch, Hugging Face Transformers, Pandas, XGBoost, YOLO, Streamlit, FastAPI
- *Tools*: Git, GitLab, DVC, MLflow, Label Studio, Docker, Jupyter, Power BI
- *Domains*: Machine Learning, Deep Learning, NLP, Computer Vision, Time Series Forecasting, MLOps, Data Visualization
- *Spoken*: Dutch (native), English (fluent)

## WORK EXPERIENCE

---

**N.V. SRO**, *Data Scientist (Part-time)*, Overijssel, Netherlands

Mar 2026 – Present

- Analyze cash register transactions and visitor data to identify trends informing operational and staffing decisions
- Audit data quality across internal systems, logging and tracking issues to ensure reporting accuracy
- Build weekly reports and dashboards in Excel and Power BI used by management for operational planning

**Scania (via Lectoraat Ambient Intelligence, Saxion)**, *Data Driven Team Intern*  
Zwolle, Netherlands

Dec 2024 – Jul 2025

- Explored operational sensor data from carrier systems to identify patterns for predictive maintenance research
- Created 50+ data visualizations using Python, Pandas, and Streamlit to communicate findings from production line data
- Built interactive Streamlit dashboards enabling the research team to explore and present sensor data findings

## PROJECTS

---

**NarcoIntel — Politieacademie**, *Computer Vision, Team of 4*

2026 – Present

- Developing a computer vision pipeline to detect IBC containers, barrels, and jerry cans in very-high-resolution satellite imagery linked to illegal chemical waste dumping
- Automated image retrieval from Satellietdataportaal and PDOK geospatial APIs
- Designed data annotation pipelines using Label Studio for consistent labeling across the team

**Kadaster AkteAI Challenge**, *NLP, Team of 3*

2025 – 2026

- Built NLP models to classify legal facts across 19,744 anonymized Dutch notarial deeds into 45 categories
- Achieved 90.2% accuracy using a 5-model extended ensemble combining Dutch BERT, TF-IDF, SVM, and Logistic Regression
- Mitigated extreme class imbalance (633:1 ratio) with stratified splits and class-weighted training

**SERENE H2020 Energy Forecasting**, *Time Series, Team of 4*

2025 – 2026

- Forecasted energy consumption for 36 households using SARIMAX, LSTM, and XGBoost models ( $R^2$  0.68–0.76)
- Deployed a full-stack dashboard with FastAPI backend and Streamlit frontend for model comparison

**Drawing Recognition AI**, *Independent Project*

2025

- Built a neural network from scratch using only NumPy for real-time sketch classification across 10 categories
- Trained the model on Google's Quick, Draw! dataset with live confidence scoring during inference