

Devika Rajasekar

Netherlands | devika2006@gmail.com | devikabuilds.pages.dev | in.devikarajasekar | devika1402

Profile

Computer science graduate with an MSc in Data Science from Leiden. At Prysmian, I built the pipelines and KPI definitions that senior teams across 53 manufacturing sites use for operational and financial decisions. On a separate project, I reconciled fragmented payment and CRM records from different systems into a unified customer view. The root cause of the mismatches was an identifier that two upstream systems referenced and neither kept updated.

Education

Leiden University, Netherlands Sep 2024 - Aug 2026
Master of Science in Computer Science: Data Science

Vellore Institute of Technology, India Sep 2020 - Jun 2024
Bachelor of Technology in Computer Science Engineering (spl. in AI & Robotics)

- CGPA: 8.5/10.0

Experience

HR Digital Data Science Intern Aug 2025 - Aug 2026
Prysmian, Delft, Netherlands

- Engineered end-to-end ETL data pipelines in Qlik Sense using modular load scripts and mapping tables, delivering trusted financial and operational KPIs to senior stakeholders across 53 manufacturing sites.
- Partnered with cross-functional stakeholders to translate business reporting needs into reusable data models, KPI definitions, and dashboard logic (variance analysis, conditional indicators).
- Ran MSc thesis research on AutoML explainability: tested whether SHAP-based feature importance explanations stayed consistent across near-optimal models (the Rashomon set) produced by AutoML systems for time series forecasting.
- Built and ran reproducible experiments using AutoGluon and H2O AutoML across six real-world benchmark datasets.

Deep Learning & Image Processing Research Intern May - Jul 2023
Centre for Cyber Physical Systems (CCPS), VIT Chennai, India

- Engineered a predictive deep learning model (AdaptUNet) for prevention of colorectal cancer; achieving segmentation accuracy to 0.9104 Dice coefficient.
- Applied cross-dataset generalisation tests and comprehensive data-validation checks; work published as first-author article [Elsevier Heliyon 2024](#).

Data Analysis Intern Jul - Nov 2023
School of Computer Science & Engineering, VIT Chennai, India

- Designed a deep neural network in R for pulsar star classification; reaching 98% accuracy, 0.8749 Kappa on the HTRU2 dataset.
- Led model reproducibility and compliance checks aligned with IEEE academic QA standards, underpinning successful conference publication [Published Findings](#).

Research Assistant Aug - Dec 2022
School of Computer Science & Engineering, VIT Chennai

- Performed an in-depth analysis of 20+ research papers, uncovering cutting-edge applications of mobile robots in crisis management; contributed to a published review impacting contemporary discourse on robotics in search and rescue operations.

Projects

Lead Conversion Data Product Design [lead-conversion-data](#)

- Architected a relational schema for lead-to-member conversion, containerised a Postgres database, and populated it with API-driven synthetic datasets simulating real-world lead progression.
- Developed optimized SQL queries to monitor revenue milestones and sales performance; designed a dashboard prototype to enable business managers to explore KPIs and inform decision-making.

Cross-System Customer Data Reconciliation [multi-system-data](#)

- Reconciled customer and financial records across Salesforce, NetSuite, and Stripe-style data starting from 12,566 supplied case rows, cleaning duplicate records, missing identifiers, mixed date formats, and mixed

currency formats.

- Built deterministic exact-key matching, avoiding fuzzy or probabilistic linkage, and kept unresolved links visible instead of forcing them: 49 NetSuite rows missing a Stripe ID, 675 unmatched Stripe transactions.

Fashion Analyzer

[fashion-analyzer](#)

- Built PySpark pipelines to analyze fashion sales trends (category, style, color, season, region).
- Produced stakeholder-ready outputs by exporting curated insights (CSV/JSON) and generating multi-panel dashboards and detailed visual reports.

NER in the CSIRO Adverse Drug Event Corpus

[ade-ner](#)

- Fine-tuned a BioBERT model incorporating Focal Loss to prioritize underrepresented classes. Achieved accuracy of 88.92% (F1: 88.04%), with notable improvements in rare entities: B-Finding (0% → 21.51% F1) and I-Disease (1.68% → 17.33% F1).
- Applied Focal Loss to address severe class imbalance, demonstrating model tuning for biased biomedical datasets.

Graph Anonymization Study: Topology Preservation vs Privacy

[privacy-by-design](#)

- Evaluated 4 anonymization methods across 5 real-world graphs; measured structural preservation (modularity shift <3%) and re-identification risks (<1%).
- Findings parallel regularization of correlation matrices under information constraints.

Energy Time-Series Forecasting for Hydropower

[energy-forecasting](#)

- Led a project forecasting hydro energy generation in the Eastern region of India using 16 months of real-time data from the National Power Portal.
- Built predictive models using SARIMA, achieving RMSE of 153.55 and MAE of 78.03; optimised models using AIC/BIC criteria.

Reinforcement Learning Benchmarks (REINFORCE → A2C)

- Implemented and benchmarked policy gradient methods on OpenAI Gym environments; conducted $5 \times 1M$ -step experiments with logging, smoothing, and comparison plots.
- Extended environment with transaction-cost penalties to simulate real-world P&L drift; evaluated final reward stability.

Skills

Programming Languages: Python, R, SQL

Machine Learning & Deep Learning: PyTorch, TensorFlow, Scikit-Learn, Transformers, Computer Vision (OpenCV, U-Net), PySpark

Data Analysis & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Power BI, Tableau, Qlik Sense

Data Engineering & Automation: dbt, Docker, Git, Apache Airflow

Cloud Platforms & Services: AWS (S3, EC2), Google Cloud Platform

Project Management & QA: Agile/Scrum, Data Validation, Requirement Gathering, Stakeholder Communication

Certifications:

AWS: Introduction to Machine Learning

Google Cloud: Big Data and ML Fundamentals

DataCamp: Data Engineer Professional Certificate (DE0019121449079)

Languages: English (Fluent), Dutch (A2 niveau)

Leadership & Engagement

CodeChef VIT Chennai

Jan 2023 - Jul 2023

General Secretary

- Spearheaded recruitment initiatives, successfully attracting 500+ applicants.
- Organized coding events, workshops, and seminars to foster a vibrant coding culture.

IEEE-Women in Engineering, VIT Chennai

May 2021 - Dec 2021

Active Member

- Advocated for women in STEM and organized professional development events, working with student and faculty groups.