

Niklas Lange

 [HerrNiklasLange](#) |  [Niklas-Lange](#) |  niklas62.35@gmail.com |  +44 7881856472

EDUCATION

2023 - 2027	Integrated Master's Degree of Computing Science at University of Strathclyde	(Expected 1:1)
Achievement:	Dean's list (2023 - 24)	
Coursework:	Information Access & Mining (Python), User & Data Modelling (SQL), Machine Learning	
2024 - 2025	Year Abroad at Michigan Technological University	
Achievement:	1 of 5 selected CS students to go abroad due to high academic performance & Dean's list (Spring 2025)	
Coursework:	Artificial Intelligence & Machine Learning (Python), Data Mining (Python)	

PROJECTS

ML System for Phishing Detection (Honours Dissertation)	Link to Code
– Quantified up to 48% accuracy degradation under distribution shift by evaluating on historical vs. newly collected data, evidencing the need for continuous retraining in production; differences validated with McNemar significance testing .	
– Built an end-to-end ML pipeline in Python classifying websites from raw URL, HTML and WHOIS metadata, comparing three scikit-learn models (Random Forest, Logistic Regression, KNN) against a three-branch PyTorch CNN.	
– Engineered automated data-collection pipelines over 60,000+ records (Common Crawl, OpenPhish, WHOIS) and 38 features, preventing leakage through strict train-only scaling and fixed seeds; published the resulting dataset on Kaggle .	
– Automated the full evaluation stage (metrics, ROC/ablation plots, feature-importance analysis) to run without manual intervention.	
Hierarchical Reinforcement Learning Fraud Detection (HRLFD)	Link to Demo
– Built and deployed an end-to-end fraud detection system combining Q-learning and decision trees in TensorFlow and PyTorch across 30 engineered features, cutting false positives by 20% against the baseline at equal recall.	
– Reported precision of 60% and recall of 99% (AUC-PR 98) on a class-imbalanced dataset, selecting the operating point against the cost of false positives rather than raw accuracy.	
– Led a team of 3 , delegating by expertise and running weekly stand-ups to deliver the project.	
Web Scraping and Visualization Project	Link to Demo
– Developed and maintained automated R and Python scraping pipelines, ethically sourcing over 1 million data points and applying tokenization and stop-word removal to analyse trends, saving 30 hours through automation.	
– Served results via a hosted web interface with geographic analysis, refining the data model over a year to cut query latency by 25%.	

WORK EXPERIENCE

Titanom Solutions - Software Engineering Intern	July - September 2026
– Delivered a bespoke content management and conference system for a German federal agency, serving 10 sub-organisations to a fixed go-live date.	
– Using Docker to containerise applications for consistent, reproducible deployment.	
Markerstudy Insurance - Junior Operations Analyst (Intern)	May - August 2024
– Built data-cleaning and transformation pipelines over 100,000+ record datasets, integrating with legacy systems to support post-merger digital transformation.	
– Applied data-science techniques in Python and delivered Power BI dashboards from Databricks/SQL , identifying process gaps that drove a 13% efficiency improvement.	
First Central Insurance - Intern, Advanced Machine Learning team	July - August 2023
– Performed data analysis on datasets with 1 million+ records using Python (Pandas, scikit-learn) to identify key predictive features.	
– Evaluated and communicated model performance (AUC, precision, recall) through data visualisation to the Advanced Modelling team.	
– Reviewed research on deep learning and reinforcement learning for pricing and risk applications; gained exposure to production ML models in risk, fraud and customer segmentation.	

SKILLS

Technical Skills	
Languages:	Python (Pandas, NumPy, scikit-learn), SQL, R, Java, Git
ML & Deep Learning:	TensorFlow, PyTorch, reinforcement learning, NLP, feature engineering
Statistics & Evaluation:	Hypothesis testing (McNemar), AUC / precision / recall / F1, ablation and feature-importance analysis, class imbalance
Operations & Deployment:	Docker (containerised deployment), Databricks
Data & Visualization:	Power BI, Tableau, exploratory data analysis
Certifications:	DataCamp Data Scientist Associate , Data Analyst Associate , Data Literacy
Soft Skills and Languages	
Languages:	English (native), German (native)
Community & Leadership:	Google Developer Student Club, University Hockey Captain, Tutoring
Last updated: August 23, 2026	