Paul Thielen

pthielen.ds@gmail.com +1 (445) 448-1249 Portfolio paul-thielen

EDUCATION

Carnegie Mellon University (CMU) – *Pittsburgh PA, USA* *Incoming Spring 2026*
Master of Science in Electrical and Computer Engineering (AI/ML Concentration)

- Specializing in AI/ML Systems with a focus on developing adaptable, resilient, and trustworthy algorithms for next-generation medical devices

Technische Universiteit Eindhoven (TUE) – *Eindhoven, Netherlands* *September 2020 – June 2024*
Bachelor of Science in Data Science

- GPA: 3.94/4.00 (**Cum Laude**)
- Minor: Cognitive Science & Artificial Intelligence

EXPERIENCE

Uber – *Amsterdam, Netherlands* *Jan 2023 – Jul 2023*
Data Scientist Intern

- Developed and implemented a **SARIMA model** to forecast sales quota attainments, enabling the sales team to proactively refine their strategies and achieve their targets.
- Applied time series analysis and predictive modeling techniques, resulting in an **18%** increase in weekly sales income by identifying high-impact sales strategies and improving decision-making
- Wrote and automated PrestoSQL queries, significantly reducing manual workload by **16 hours** per week with optimized query execution and report generation

RESEARCH PUBLICATIONS

Medical Data Under Shadow Attacks via Hybrid Model Inversion, *AISTATS*, *2025* | *Link*

- Developed MEDUSA, a hybrid **model inversion** framework for gray-box settings that combines gradient-based optimization with inverse modeling, achieving up to a **12%** improvement in reconstruction fidelity
- Demonstrated the use of reconstructed medical images as synthetic training samples, achieving training performance within **5%** of real data benchmarks
- Invented and evaluated the KNN Smearing Defense method, reducing reconstruction quality by up to **64%**

PROJECTS

Predicting Crimes in Barnet, London

- Led a project as **Scrum Master** to implement OLS, Random Forest, SARIMA, and Prophet models to forecast burglary trends in Barnet, training a model for each Lower Super Output Area (LSOA) to account for regional variations.
- **Prophet** models demonstrated the highest accuracy, achieving an average R^2 score of 0.83, effectively capturing **83%** of the variability in burglary rates.
- Implemented Gurobi **Integer Linear Programming** (ILP) solver to optimize officer allocation across LSOAs

Food Insecurity in South Sudan

- Led a team project in collaboration with the Zero Hunger Lab to predict food insecurity risks in South Sudan using text data from news sources and climate data, earning a **9/10 grade**
- Trained linear regression models using climate data and conflict metric, **GPT-3**-extracted mentions of violence from local and global news articles about South Sudan, to forecast monthly risk scores
- Modeled the impact of the Bentiu takeover (major conflict event) using a **Bayesian Structural Time Series Model**, revealing a **71%** (statistically significant) rise in risk scores and a lasting shift in data distributions, demonstrating the enduring impact a major conflict event has on food insecurity

SKILLS

Programming Languages	Python, R, SQL, MATLAB
Deep Learning Frameworks	PyTorch, TensorFlow, Keras, Hugging Face, JAX
Machine Learning & Data Science Tools	Sklearn, OpenCV, Numpy, Pandas, Spark, Seaborn, PowerBI, Tableu
Cloud Computing Technologies	Microsoft Azure, Amazon Web Services, Google Cloud Platform

CERTIFICATES

Machine Learning Fundamentals Track (DataCamp)
Deep Learning Track (DataCamp)
Microsoft Excel (Coursera)
Microsoft AI & ML Engineering (in progress Coursera)