(408) 375-5967
Santa Clara, CA
paetausoren@gmail.com

Soren Paetau

Data Scientist / Mathematician

Portfolio: sorenpaetau.com
github.com/sorenpgithub
linkedin.com/in/soren-paetau

A mathematician with extensive project and theoretical experience in machine learning applications to real-world classification problems. Aiming to work under a data heavy role where I can use my problem solving skills and communication abilities.

Education

California Polytechnic University - San Luis Obispo Spring 2024
Bachelor of Science in Mathematics GPA: 3.52^*
Cross Disciplinary Studies Minor in Data Science, *20+ classes and Extensive Project Experience*
Resident Advisor, University Housing 2021 — 2024
Cal Grant Recipient

Technical Experience

Boiling Classification with CNN **Jan 2024 — Present**
NASA Ames Research Center *Palo Alto, CA*

- Performed analysis of performance on convolutional neural network on audio classification problem in MATLAB
- Educate and report relevant in audio classification through comprehensive review of relevant literature

Research Assistant **Jan 2024 — Present**
California Polytechnic University - San Luis Obispo

- Conducted comprehensive reviews of computer vision literature, synthesizing complex mathematical concepts and incorporating them into the theoretical foundations
- Formulated and proved theorems concerning the essential matrix in mutli-spatial computer vision applications

Fellow / Research Assistant **Aug 2022 — Jun 2023**
Central Coast Data Science *San Luis Obispo, CA*

- Translated and implemented existing library of 1000+ lines of R code into a ranking problem library in Python
- Produced theories and ideas concerning weighted rank aggregation, attending graduate-level meetings
- Participated in weekly seminars, discussing topics centered around data science and statistical learning

Spotify User Analysis **Nov 2023**
California Polytechnic University - San Luis Obispo

- Preprocessed data, extracted features, and applied advanced analytical methods to derive meaningful insights in Python using *numpy* and *scikit-learn*
- Publicly presented results and typed up a formal report in LaTeX

Current **Jan 2022**
Camp Poly Hacks *San Luis Obispo, CA*

- When presented with USGS flood data, pitched an app Current, utilizing a theoretical machine learning model to predict floods on future dates
- Responsible for developing the technical overview, referencing existing research, and creating the presentation
- Answered judge's questions concerning the wealth of data and details of our proposed ML model

Internship **Aug 2019 — Jun 2020**
Q Pi Education *Santa Clara, CA*

- Created a competitor to an educational autonomous vehicle starter kit for young STEM students to affordably and easily begin their journey into the application of CAD and Python programming.
- Developed and pitched software while participating in the full lifecycle of the program.

Skills

Languages	MySQL, R, Java, Git, LaTeX, Python *(highly proficient)*
Libraries	*pandas*, *NumPy*, *tidyverse*, *scikit-learn*, *TensorFlow*
Presentation Tools	Technical Writing, Public Speaking, Agile Communication/Development

Coursework

Pure Mathematics *with focus in Proof Writing and Theory* 2020-2024
Applied Linear Models and Multivariate Statistics
Database and Distributed Computing Systems
Analysis of Algorithms and Knowledge Discovery from Data
Advanced Statistical Computing and Visualization with R