Samu Syrjänen | University of Helsinki / Aalto University

samu.syrjanen@gmail.com
+358 404161217
My Website
Location: Helsinki, Finland
Willing to relocate globally

Field:
Data Science
Data Engineering
Data Analysis
Machine Learning

Languages:
English (CEFR C1)
Finnish (Native)
Japanese (Beginner)

About Me

I'm a **Data Science Master's student** at the University of Helsinki, and **research assistant** at Aalto University with one year of experience. My background is in **Computer Science**, **Machine Learning**, **Data Engineering**, and **Data Analysis**. I'm looking for long-term work opportunities to gain experience and develop more specialized skills.

Future career interests include working with data pipelines, analytics, cloud platforms, and machine learning algorithms to provide solutions for **product development**, **marketing**, and **business intelligence** problems. Besides the more technical roles, I'm also able to work in **business administration** where a more tech-heavy background might be beneficial.

Skills

- Python
- SQL
- Many Kinds of ML Algorithms, Including Convolutional Neural Networks
- ETL/ELT Pipelines
- Spark
- Kafka
- PowerBI
- PyTorch
- AWS
- Databricks
- Data Cleaning/Tests
- Communication and Coordination
- Scrum and Agile Development

Experience

Dec 2024 - Current

Research Assistant: Hyperspectral Image Processing - Aalto University

As part of Jaan Praks research group at Aalto University and ESA's Hera space mission, I'm responsible for creating a pipeline that transforms hyperspectral images from space into final data products.

- Built a modular multi-level pipeline that mainly uses **Python** (**Numpy**, **Pandas**, **OpenCV**, **SciPy**, and **Matplotlib**) to calibrate the incoming hyperspectral images based on measured calibration statistics and metadata.
- Extracted the positions, orientations, and other geometric information of the target asteroids and camera with **SpiceyPy** Python library and ESA's **SPICE kernels**, based on data acquisition time.
- Set up a pipeline that estimates the mineral composition of the calibrated asteroid spectra with a **Convolutional Neural Network** ML model.
- Actively **coordinated tasks** and **requirements** between cross-national teams working on this project.
- Made **Planetary Data System** (**PDS4**) products for archiving the produced data into ESA's Planetary Science Archive.

Sep 2023 - Current (Expected Early 2026)

Master's Degree in Data Science - University of Helsinki | Transcript of Records

May 2024 - Aug 2024

Research Assistant: ML Model for Hyperspectral Data - University of Helsinki | Certificate

- Created a **Convolutional Neural Network** enhanced **Gaussian Process** algorithm for estimating asteroid surface age based on hyperspectral reflectance measurements.
- Proved that the GP algorithm is surprisingly flexible even with a **sparse training set** of 169 entries.
- Surpassed the performance of a competing ensemble model with R^2 of 0.9934 vs 0.9905.
- Coauthored a scientific paper comparing the GP and ensemble models. (Published as a journal article soon!)

Sep 2019 - Dec 2023

Bachelor's Degree in Computer Science - University of Helsinki

Thesis (Not finished yet)

Scalable Data Streaming Pipeline in Cloud Environment

- Developed an end-to-end streaming ELT pipeline with **Kafka**, **Spark/Databricks**, and **PowerBI**.
- Transformed real-time stock market trades into an aggregated candlestick (OHLC) **time-series** dataset and derived **analytics** from it.
- Used **Lakehouse**, **Medallion Architecture**.
- Built all components on scalable **cloud compute**.

My other projects/products include:

1. [Collab] [ML] Building Façade Recognition
2. [Collab] [Agile] Mobile App Development
3. [SQL] Database Project: Forum Website
4. [Collab] [ML] Exploratory ML Project
5. [ML] K-Means Clustering for Text Data

See my website for details...