



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

Field:  
Data Engineering  
Data Science  
Machine Learning  
Data Analysis

Personality: INTJ-A  
Languages:  
English (CEFR C1)  
Finnish (Native)  
Japanese (Beginner)

## About Me

I'm a **Data Science Master's student** at the University of Helsinki, and a **research assistant** at Aalto University. My background is in **Computer Science, Data Engineering, Machine Learning, and Data Analysis**. I have also studied Physics and Mathematics, and know a thing or two about sailboats and Leopard 2A6 tanks, and how to lead their crews.

I'm looking for long-term work opportunities to gain experience and develop more specialized skills. Future career interests include working with **cloud platforms, data pipelines, analytics, and machine learning** to provide solutions for product development, marketing, and business intelligence problems. Besides the technical roles, I'm also able to work in **business administration** or as a **product manager**, where a more tech-heavy background might sometimes be beneficial.

## Skills

- Python
- SQL
- Excel
- Many kinds of ML algorithms (e.g. CNN)
- ETL/ELT pipelines
- Spark
- Kafka
- Grafana
- PyTorch
- AWS
- Databricks
- Team leading
- International coordination
- Scrum and agile development

## Experience

Sep 2023 - Current (Expected in H1 2026)

**Master's Degree in Data Science** | University of Helsinki | [Transcript of Records](#)

Dec 2024 - Nov 2025

**Data Engineer: Hyperspectral Image Processing** | Aalto University ([ESA's Hera Space Mission](#)) | [Certificate](#)

- 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** 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

May 2024 - Aug 2024

**ML Engineer: GP Model for Hyperspectral Images** | University of Helsinki | [Journal Article](#), [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 measurements
- Surpassed the performance of a competing ensemble model with  $R^2$  of 0.9934 vs 0.9905
- Co-authored a scientific journal article comparing the GP and ensemble models ([Journal Article](#))

Sep 2019 - Dec 2023

**Bachelor's Degree in Computer Science** | University of Helsinki

Jul 2020 - Jun 2021

**Non-Commissioned Officer and Leopard 2A6 Commander** | The Finnish Defence Forces

## Master's Thesis

### 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
- Implemented a **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 some details...