Samu Syrjänen | Student at the University of Helsinki

samu.syrjanen@gmail.com
+358 404161217
Linkedin
Helsinki

Field:
Data Science
Machine Learning
Computer Science

Languages:
English
Finnish
Swedish (weak)

About Me

I have experience in software engineering, data science, and machine learning from 229 cr worth of university courses. My overall average grade is 3.76/5. In software engineering, data science, and ML courses, I have an average grade of 4.18/5. I have worked as part of a group in a total of 6 different software or data-related projects, and I have learned that communication and help between group members go a long way. I value learning opportunities and the continual growth of skills. After graduation, I envision myself working as a data scientist or data engineer in a team-oriented role.

Education

Sep 2023 - Ongoing

Master's Degree in Data Science – *University of Helsinki*
Graduation expected in: Autumn 2025

Sep 2019 - Dec 2023

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

Skills

- Python
- SQL
- Scrum and agile development
- Continuous integration
- Frontend & Backend
- End-to-end machine learning (MLOps) with Kubernetes
- Data analysis
- JavaScript/TypeScript
- R (weak)

Experience

Oct 2023 - Dec 2023

Introduction to Machine Learning: Group Project – *University course*
https://samusyrjanen.github.io/CV_Samu_Syrjanen/IML_project_report

In this project, our task was to predict the saturation vapor pressure of different molecules based on their features. To achieve this, we experimented with multiple different machine learning models and different techniques to find and train the best possible model for the task. The project report contains a demonstration of exploratory data analysis, principal component analysis, different ML models, hyperparameter optimization, feature selection, and our conclusion.

There was a Kaggle competition regarding the project (link to the competition), where each group would post their results. Our group performed averagely. Our group's R-squared score was 0.6729 while the winning group had a score of 0.7275.

Essential skills:

- Exploratory data analysis (EDA)
- Feature selection
- Decision trees
- Random forest

- Hyperparameter optimization
- Data processing and handling
- Linear regression

- Gradient boosting
- Python
- Documentation

Jan 2023 - May 2023

Software Engineering Group Project: Berry Picker Tracker – *University course*
https://marjanpoimijat.github.io/

A phone application which tracks users and prevents them from getting lost in a forest. Users can see the routes of other tracked users. If internet connection is lost, the latest location can be seen by others. More extensive demonstration on our github page.

The course's aim was to simulate a real job. Our group resumed the work of a previous group. We experienced many problems regarding the compatibility of dependencies and other technologies used, such as React Native, Expo, map tiles, and caching.

Our biggest achievements:

- Deploying the app to the university server
- A working testing version on a real phone (not just emulator)
- Code refactoring: code is easily readable and maintainable (quality)
- A setup script that enables a fast start for a new developer to begin development
- Updating documentation
- Updating dependencies
- All user data is encrypted and safe
- Ability to add, track, and discard multiple other users at the same time (+ settings and UI for it)
- The map tile source can be changed
- Ability to switch languages
- UI overhaul
- Other quality of life features

Essential skills:

- Full stack development
- Python (for backend)
- TypeScript (for frontend)
- Group leading
- Understanding customer needs
- Scrum and agile development
- Development in a group
- Testing
- Software architecture
- Documentation

May 2023 - June 2023

Text Clustering Solo Project: Sortext – *University course*
https://github.com/samusyrjanen/sortext

This app clusters similar texts into groups. It is self-made from the ground up, using no premade algorithms from external libraries. The app uses a dataset of old BBC news articles and takes additional texts as input from the user. It sorts those articles into groups based on text similarity using TF-IDF matrix and k-means clustering algorithm.

Essential skills:

- Python
- Unsupervised Machine Learning (K-means with TF-IDF matrices)
- Data processing and handling
- Testing
- Software architecture
- Documentation

See more experiences…