https://github.com/sonalexle
https://www.linkedin.com/in/sonalexle/

Son Le

+41 79 506 75 17
minhsonle199@gmail.com

Experience

Research Assistant — **Aalto University** — **Jan 2022 – Aug 2022**

Computational Systems Biology group (Jun 2022 – Aug 2022)

- Helped build a novel latent-variable deep neural network (NN) for synthetic data generation and causal effect inference on heterogeneous data from both observational and experimental sources.
- Conducted experiments on local cluster. The proposed model showed improvement over key NN baselines.

Probabilistic Machine Learning group (Jan 2022 – May 2022)

- Researched the applicability of Bayesian projection predictive inference (projpred) for causal analysis.
- Showed that projpred has potential but needs modification for causal analysis.

Machine Learning Trainee — **Nokia** — **May 2021 – November 2021**

- Evaluated causal inference by taking initiative to research and implement relevant concepts and tools, leading to successful presentations to colleagues and business partners.
- Trained in basic software development practices. Presented math concepts to software engineers.

Teaching Assistant — **Aalto University** — **March 2021 – May 2021**

- Course: Programming 2. Assisting students in solving programming tasks in Scala.

Education

Switzerland — **EPFL** — **Sep 2022 – Present**

- M.Sc. in Data Science
- Courses: Applied Data Analysis, Machine Learning, Statistics for Data Science

Finland — **Aalto University** — **Sep 2019 – May 2022**

- B.Sc. in Data Science. GPA: 5 out of 5
- Courses: Bayesian Data Analysis, Programming Parallel Computers, Data Structures and Algorithms

Technical Experience

School Projects

- **Viral tweets MLOps**. Pipeline for BERT large language model predicting tweet virality. Integrated model with data collection, retraining, and deployment. Python, PyTorch, MLFlow.
- **GeekLearning**. Dockerized web app for classification of cartoon images. Implemented SVM, feedforward net, and a state-of-the-art neural net. Flask, PyTorch, Docker.
- **Energy Demand Prediction**. Modeled a complex time series on the electricity demand in Helsinki. Used SARIMAX, recurrent neural net, and Facebook Prophet. Python, TensorFlow, Prophet.
- **BayesianRegression**. Bayesian workflow for predicting electrical output of a power plant. Bayesian linear and generalized additive models. R, Stan, brms.
- **RegressionApp**. Fits and visualizes univariate linear and quadratic regression models. Fits models on gigabytes-sized datasets in a few minutes via online model updating. Scala, Java.

Additional Experience and Honors

- **Data Science Tutor, 2020 – 2021** Assisting new Aalto students in adapting to campus life.
- **100% tuition waiver** Merit scholarship for non-EU/EEA students, awarded for the whole degree program.
- **Dean's Incentive Scholarship, 2020 – 2021** Students meeting the annual target of 60 credits.

Updated September 29, 2022