Mikolaj Mazurczyk

mikmaz.github.io
mima@di.ku.dk
mikmaz
mmazurczyk

Research interests:
conformal prediction, PAC-Bayesian analysis, remote sensing, probabilistic machine learning, computer vision, reinforcement learning

Education

University of Copenhagen 10/2024 – present
Ph.D. in Machine Learning Copenhagen, Denmark
Affiliated with the Pioneer Centre for AI

University of Copenhagen 09/2021 – 09/2023
MSc in Computer Science Copenhagen, Denmark

University of Warsaw 09/2019 – 06/2021
BSc in Computer Science Warsaw, Poland

University of Wroclaw 09/2018 – 06/2019
BSc in Computer Science Wroclaw, Poland
Finished the first year of my BSc here, and for the following two years, I transferred to the University of Warsaw

Work Experience

Research Assistant 11/2023 – 08/2024
University of Copenhagen Copenhagen, Denmark

- Was hired at the DARK, an astrophysics research section of Niels Bohr Institute, to work towards publishing the results of my Master's thesis
- Developed a deep learning method for predicting the properties of the cosmic dust produced by supernovae based on the data collected with the James Webb Space Telescope (JWST)
- Worked within probabilistic machine learning, which involved training a neural network predicting a multivariate normal distribution, evaluating its calibration, and Markov Chain Monte Carlo sampling (more details under Projects)
- Towards the end of my employment, I worked on a project utilizing statistical models and dimensionality reduction methods to improve Type Ia supernovae standardization to help solve the Hubble constant tension

Teaching Assistant 09/2022 – 04/2023
University of Copenhagen Copenhagen, Denmark

- Worked part-time as a teaching assistant in three subsequent 7.5 ECTS Master's level courses: Machine Learning A (MLA), Machine Learning B, and Online and Reinforcement Learning (OReL).
- Primary responsibilities included holding three-hour Q&A sessions, where I helped students with the weekly assignments and whole course material, and handling grading
- Additionally, I helped in developing the exam for MLA and recorded a guide to a reference solution for a coding assignment in OReL

Private math tutoring 09/2018 – 06/2022

Throughout most of my academic education, I worked as a private math tutor for high school and primary school students. Witnessing the progress among young students in understanding mathematics, who had lost confidence in their abilities, stands as one of the most fulfilling work-related experiences I've encountered.

Student Software Engineer 10/2020 – 06/2021
ScyllaDB Warsaw, Poland

- As a part of our Bachelor's thesis, I developed a Kubernetes Cluster Autoscaler for Scylla with three other students from my program
- Participated in every step, from writing the architecture proposal to implementing, testing, and producing documentation
- The project can be seen at github.com/scylladb/scylla-cluster-autoscaler

Data Analysis Intern 09/2019 – 10/2019
SophScope Wroclaw, Poland

- Created reports for marketing purposes based on usage metrics of the *Kiedy Wykład* mobile application that allows students to view class schedules and grades for selected Polish universities
- Analyzed database content, wrote SQL queries for adequate data aggregation, and created reports in PowerBI based on the queried data

PROJECTS

Estimating cosmic dust properties using Bayesian inference | ***Pytorch*** 02/2022 – 09/2023
University of Copenhagen Copenhagen, Denmark

- Worked on the project applying deep learning to the task in astrophysics in terms of 30 ECTS Master's thesis under the supervision of Prof. Oswin Krause, with Prof. Christa Gall as a co-supervisor
- The goal was to derive an accurate representation of the intricate distribution of the cosmic dust features ($\boldsymbol{y}$) conditioned on observations of the JWST ($\boldsymbol{x}$)
- As a proposed solution, I trained the deep learning model representing more tractable density $p(\boldsymbol{x}|\boldsymbol{y})$, and then by choosing suitable prior $p(\boldsymbol{y})$, I retrieved the posterior by sampling from $p(\boldsymbol{y}|\boldsymbol{x}) \propto p(\boldsymbol{x}|\boldsymbol{y})p(\boldsymbol{y})$ with MCMC
- Additionally, I compared the effectiveness of a residual neural network (ResNet) and a more complex Conditional Variational Autoencoder (CVAE) in modeling $p(\boldsymbol{x}|\boldsymbol{y})$
- The approach gave promising results, likely competitive with the standard methods, where CVAE only slightly outperformed the ResNet
- The implementation and the thesis can be found at github.com/mikmaz/jwst-space-dust-thesis.

Hierarchical Normalizing Flow for ECG data | ***Pytorch*** 09/2022 – 01/2023
University of Copenhagen Copenhagen, Denmark

- Worked on the project in terms of a 15 ECTS Project Outside the Course Scope under the supervision of Prof. Oswin Krause, in cooperation with the Department of Biomedical Sciences, with Prof. Jørgen K. Kanters and Dr. Jonas L. Isaksen as the advisors
- Successfully designed, trained, and evaluated a Hierarchical Normalizing Flow Network applied to the electrocardiographic (ECG) data to learn the invertible mappings of patients' features to their ECG
- Unfortunately, the quality of samples was insufficient to follow with the final analysis of how patients' attributes are mapped in the distribution learned by the model
- The implementation and the final report can be found at github.com/mikmaz/ecg-norm-flow

SKILLS

Programming:	Python (`PyTorch`, `NumPy`, `(geo)pandas`, `Astropy`, `Pyro`), Haskell, Golang, C, C++, Java
Software:	LaTeX, Slurm, QGIS, Kubernetes
Language:	Polish (native), English (C1, CAE grade B)