Jakub Wisniewski

+45 31 78 52 46 | jakwisn@gmail.com | linkedin.com/in/jakwisn | github.com/jakwisn | jakwisn.github.io

Skilled data scientist with expertise in AI and experience in researching and developing DS and ML solutions. Demonstrated leadership through academic and professional contributions, including research, mentorship, and organizing events. Upon graduation, I intend to leverage and further advance my skills in AI.

EDUCATION

Danmarks Tekniske Universitet Denmark, Copenhagen
Master in Artificial Intelligence *Feb. 2023 - ongoing (estimated end Feb. 2025)*

Warsaw University of Technology Poland, Warsaw
Bachelor of Science in Data Science *Oct. 2018 – Feb. 2022*

TECHNICAL SKILLS AND LANGUAGES

Programming: Python, R, SQL (MySQL), JavaScript, HTML/CSS
Frameworks and Libraries: Numpy, Pandas, PyTorch, Tensorflow, FastAPI, Plotly, Scikit-learn, Langchain, OpenCV, Mlflow, Weights&Biases, Dash, Selenium,
Other: Git, Docker, Linux, LaTex, HPC, Slurm, Tableau, PowerBI, GCP, VS Code, PyCharm, NoSQL (Neo4j)
Languages: English - Proficient (CAE on level C2), Polish - Native

EXPERIENCE

Data Scientist Feb. 2022 – Present
SEB *Copenhagen, Denmark*

- Developed payment data processing pipeline using SQL, Pandas, Numpy, and Scikit-learn that fed data into Neo4j database, which was used as a backend for a client-facing analytics tool.
- Created containerized applications with Docker, SQLite, and FastAPI hosted via Cloud Run on Google Cloud.
- Trained time-series Prophet model, where I designed and tracked experiments using MLflow. The model was served on GCP via CI-CD pipelines from hooks on a GitHub repository. API with model predictions was then integrated into the client-facing dashboard.
- Created a web scraper using Selenium and BeautifulSoup to systematically gather data from different sources and automatically upload it into Sharepoint.
- Created a set of template notebooks that were used for generating interactive HTML reports using clients' data and visualizing it with Plotly and Highcharts
- Applied a wide range of machine learning models to different problems like multiclass classification and anomaly detection using auto-encoders created in PyTorch.
- Generated a series of annual report summaries using LLM and NoSQL database containing 30GB of text data
- Conducted client presentations using tools developed by our team.

Research Software Engineer March. 2020 – Feb 2022
MI2 DataLab *Warsaw, Poland*

- Developed and maintained R and Python open-source packages for Explainable Machine Learning and Fairness (fairmodels and DALEX).
- Created a large-scale database of medical DICOM images that researchers, engineers, and radiologists used.
- Trained PyTorch deep learning models on a multi-GPU cluster within a Slurm environment.
- Created a tailored text annotation tool in Dash in close collaboration with radiologists that sped up the annotation process multiple times.
- Published articles, blogs and designed landing pages for the software.
- Conducted presentations (f.e. regular talk and workshops on useR! 2021) and presented team's work on poster session on CVPR workshop.

LEADERSHIP AND OTHER ACTIVITIES

- President of Data Science Club - coordinated projects with members about the use of AI in archeological scans, took part in Kaggle competitions, hosted talks with industry leaders as well as hackathons and events promoting data science in the academic community (Like PoweR - data science workshops for beginners in R and Python).
- Co-organized Hackathon4Ukraine where participants had a chance to create valuable solutions to help Ukrainian people. We managed to raise 33 000 DKK for Polish Humanitarian Action.
- Conducted a commercial Python course tailored to the needs of PhD biology students. The course covered fundamental language concepts, data manipulation, visualization, and machine learning.
- Presented and won the award for the best poster at the MLinPL conference in 2020.
- Player and member of Lyngby Tennis Club