

# Mateusz Jurewicz

Senior Machine Learning Engineer






Ph.D. in Computer Science, IT University of Copenhagen

[mateusz.jurewicz@gmail.com](mailto:mateusz.jurewicz@gmail.com)



## Areas of specialization

• artificial intelligence • natural language processing  
• AI safety • structure prediction • large language models • generative AI • neural clustering • ML ops

 [github.com/mateuszjurewicz](https://github.com/mateuszjurewicz)  
 [in/mateusz-jurewicz-a084aa6a](https://in/mateusz-jurewicz-a084aa6a)  
 [orcid.org/0000-0003-4211-2185](https://orcid.org/0000-0003-4211-2185)  
 [scholar.google.com](https://scholar.google.com)  
 [mateuszjurewicz.wordpress.com](https://mateuszjurewicz.wordpress.com)

## Programming Skills

**Machine Learning / AI** | • PyTorch • TensorFlow • Keras • Fastai • scikit-learn

**ML Ops** | • MLFlow • Neptune • Ray • Kubernetes • AnyScale • Weights & Biases

**Cloud** | • Google Cloud Platform (GCP) + Vertex AI • Amazon Web Services (AWS)

**Scientific Computing** | • NumPy • pandas • Jupyter • Matlab • Octave

**Scripting Languages** | • Python • Golang • JavaScript • shell

**Databases** | • BigQuery • PostgreSQL • MySQL • MongoDB • GraphQL

**Other** | • ElasticSearch • Tableau • Citus • VBA

## Professional experience

2017 - present

TJEK A/S

Senior Machine Learning Developer

*Software Engineering Department*

Senior ML Developer

2022 - present

Industrial Ph.D. Candidate

2019 - 2023

Machine Learning Engineer

2018 - 2022

Machine Learning Internship

2017 - 2018

Design & implementation of machine learning & deep learning solutions in natural language understanding (NLP/NLU), image content analysis and information retrieval in **Python** and **Golang**. Training and fine-tuning LLM-based multi-label classifiers on multi-modal data to an accuracy of over 97.8% and an F1 score of over 0.89, currently in production. Elements of data engineering (BigQuery optimization) and visualization (Data Studio, Tableau). Agile development, deployment and monitoring of services in both AWS (Kubernetes) and Google Cloud Platform.

Successful self-started industrial Ph.D. grant application followed by combining work within data engineering, analytics and ML with academic research in set-input deep learning, implicit clustering and permutation learning through transformer neural networks. Publications in top venues (ICML, NeurIPS, IJCAI-ECAI, JAIR).

2017

INTEL

Software Engineer

*Nervana Systems*

Development of the flagship Nervana deep learning platform framework in Python. Close cooperation with the US-based team with regards to data curation, comprehensive testing and the implementation of inference job scheduling for the ML tool developed primarily for academic and industrial research centres.

2015 - 2017

WNS GLOBAL SERVICES

L&D Executive

*Learning & Development*

L&D Associate

2015 - 2016

Overseeing ROI projects (+\$2,500 savings) and VBA automation. Designing and delivering L&D frameworks. Heavy involvement in mentoring and knowledge-sharing programs, developing partnerships with local universities through career guidance projects and LEAN / Six Sigma courses.

2014 - 2015

MINDPOWER LTD.

Business Analyst

*IT Department*

In-depth analysis of the IT healthcare market potential for a number of EU countries. Recommendation with regards to entry strategy & risk assessment. Statistical modelling & visualizations for internal and external use. Aiding in project coordination with private and government organizations.

2013 - 2014

CHECKWARE AS

Assessment Developer

*IT Department*

Implementing online versions of scientifically supported psychometric tools and other Health 2.0 solutions, assisting in client consulting. Elements of coding in HTML, CSS and JavaScript.

## Education

2019 - 2023

### PhD, Computer Science

*Industrial PhD Program (IFD), Department of Computer Science*

IT UNIVERSITY OF COPENHAGEN, DENMARK

ADVISOR: Leon Derczynski, ITU | COMMITTEE: SEBASTIAN RISI, OLE WINTHER, NAVDEEP JAITLEY

2023

### Visiting Researcher

*Vector Institute for Artificial Intelligence, Research Stay Abroad*

UNIVERSITY OF GUELPH, CANADA

2009 - 2014

### MSc, Cognitive Science (magna cum laude)

*Rector's Scholarship for Highest Academic Achievement | 3x Laureate*

UNIVERSITY OF WARSAW

ADVISOR: Mirosław Kofta, UW; Anna-Szuster Kowalewicz, UW

2012 - 2013

### Erasmus Scholarship, Istanbul

*Department of Psychology, Faculty of Applied Sciences*

DOĞUŞ UNIVERSITESİ

## Languages

*English* C2 | CPE British Council

*Polish* C2 | native speaker

*Danish* A2+ | Danskuddannelser DU3.2

*German* A2

*Turkish* A1

## Publications

### 1. Conference papers

**M. Jurewicz**, G. Taylor, L. Derczynski (2023) The Catalog Problem: Clustering and Ordering Variable-Sized Sets. *Published in the proceedings of the Fortieth International Conference on Machine Learning (ICML-2023)*. [\[pdf\]](#)

**M. Jurewicz**, L. Derczynski (2022) Set Interdependence Transformer: Set-to-Sequence Neural Networks for Permutation Learning and Structure Prediction. *Published in the proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-ECAI 2022)*. [\[pdf\]](#)

**M. Jurewicz**, L. Derczynski (2021) PROCAT: Product Catalogue Dataset for Implicit Clustering, Permutation Learning and Structure Prediction. *Published in the NeurIPS 2021 Dataset and Benchmark Track*. [\[pdf\]](#) [\[doi\]](#)

### 2. Journal articles

**M. Jurewicz**, L. Derczynski (2021) Set-to-Sequence Methods in Machine Learning: a Review. *Published in the Journal of Artificial Intelligence Research* 72. [\[pdf\]](#)

### 3. Datasets

**M. Jurewicz**, L. Derczynski (2021) Product catalogue dataset of over 1.5 million product offers, composed into complementary sections for visual presentation. *Figshare*. [\[html\]](#) [\[github\]](#)

## Selected projects

- 2019 Google Brain's Tensorflow Speech Recognition Challenge (Kaggle Entry). [\[github\]](#)
- 2018 Visual Style Transfer as Introduction to Machine Learning for the Børnhack Summit. [\[github\]](#)
- 2017 Tic-Tac-AI. Play against and look into the mind of a tic-tac-toe player via a Django web app. [\[app\]](#) [\[html\]](#) [\[github\]](#)

## Selected Press

- 2019 [Bornhack: Kunstig intelligens har også mørke sider.](#) *Akademikerbladet*
- 2018 [Natural Language Preprocessing for Multilabel Classification.](#) *Medium*

## References

- [Sarah Vang Nøhr](#). *Engineering Manager at Tjek (2018-2020)*.
- [Professor Leon Derczynski](#). *Primary advisor during PhD Project at the IT University of Copenhagen (2019-2023)*.
- [Professor Graham Taylor](#). *Research supervisor during the stay abroad at the Vector Institute in Canada (2022)*.