Nikita Kozodoi, Ph.D.

ML scientist working on the frontier of research and business. With Ph.D. in ML and certified AWS expertise, practical experience in diverse ML/DL areas and strong communication skills, I am passionate about using AI to solve business problems.

🌐 kozodoi.me (portfolio & ML blog)   in linkedin.com/in/kozodoi   github.com/kozodoi   ✉ n.kozodoi@icloud.com

Experience

07/2022 – present **Applied ML Scientist**, *Amazon Web Services*, Berlin.
- developing cutting-edge ML/DL solutions to solve diverse business problems across industries
- interacting with large-scale customers to identify and implement high-value ML use cases
- leveraging AWS infrastructure to design, build and evaluate custom ML solutions

04/2022 – 06/2022 **Project-Based Data Scientist**, *Simply Rational*, Berlin, freelance.
- identified customer segments and prototyped segment prediction models from incomplete data
- analyzed A/B tests results to determine treatments improving customer repayment by up to 6%

04/2018 – 03/2022 **Data Science Research Associate**, *Monedo / Humboldt University*, Berlin.
- developed sample bias mitigation algorithms to improve model profitability by up to 3%
- reduced data acquisition costs by up to 25% using multi-objective feature selection methods

Education

04/2018 – 06/2022 **Ph.D. in Machine Learning, Information Systems**, *Humboldt University*, Berlin.
Published 5 papers on ML for credit risk analytics with over 100 total citations. [Papers]

10/2015 – 12/2017 **M.Sc. in Economics, Quantitative Methods**, *Humboldt University*, Berlin, GPA: 1.30 (A).
Focused on Data Science and Predictive Analytics. Thesis topic: ML for box office forecasting.

09/2010 – 06/2014 **B.Sc. in Economics**, *Higher School of Economics*, Moscow, GPA: 1.31 (A).
Focused on Statistics and Econometrics. Top-1 in the 4-year student ranking. Diploma with Honors.

Certifications

07/2022 **AWS Certified Cloud Practitioner**, *AWS*. Demonstrated foundational knowledge of AWS cloud services.
01/2022 **AWS Certified Machine Learning**, *AWS*. Learned training, tuning and deploying ML models on AWS.
06/2020 **Deep Learning Nanodegree**, *Udacity*. Implemented CNNs, RNNs and LSTMs in PyTorch.
05/2020 **Machine Learning Engineer Nanodegree**, *Udacity*. Deployed ML models on Amazon SageMaker.

Selected ML Projects

05/2021 – 08/2021 **Text Readability Prediction.** Developed a PyTorch pipeline for text complexity prediction with transformers. Deployed a web app for scoring custom texts. Placed top-9% in the Kaggle competition. [App] [GitHub]

03/2021 – 06/2021 **Image-to-Text Molecular Translation.** Translated molecule images into chemical formulas with PyTorch. Built a CNN-LSTM architecture for image captioning. Placed top-5% in the Kaggle competition. [GitHub]

09/2019 – 04/2021 **Fairness: Package for Fair ML.** Building and maintaining an R package with over 17k total downloads. Implemented calculation and visualization of fairness metrics. Wrote documentation. [Blog post] [GitHub]

09/2018 – present **Kaggle Master.** Won 18 competition medals. Top-1% of the global user ranking. [Projects]

Selected ML Publications

06/2020 – present **ML Blog.** Publishing ML/DL tutorials and project overviews. Reached over 10k monthly views. [Blog]
11/2021 **Fairness in credit scoring: Assessment, implementation and profit implications**, *EJOR*. [Paper] [GitHub]
04/2020 **Shallow self-learning for reject inference in credit scoring**, *ECML PKDD*. [Paper] [Meetup talk]
05/2019 **A multi-objective approach for profit-driven feature selection in credit scoring**, *DSS*. [Paper]

Technical Stack

Tools </> Python, R, SQL *(interm)*   🖥 VSCode, Jupyter   ☁ AWS (SageMaker, EC2, Lambda)   👥 Git, Jira

Packages ▦ Pandas, NumPy   PyTorch, TensorFlow *(interm)*   Scikit-learn, XGBoost   Huggingface, Timm

Soft Skills

Languages **English** (C2), **German** (C1), **Russian** (native)

Key skills
- strong motivation to learn and improve
- attention to detail and scientific mindset
- excellent communication skills [Public talks]
- rich experience of working in team projects