

Nikita Kozodoi, PhD

ML scientist working on the frontier of research and business. With PhD in ML and certified AWS expertise, rich experience in generative AI and diverse areas of ML, strong communication and public talking skills, I am enthusiastic about using AI to solve challenging business problems across industries.

kozodoi.me (portfolio & ML blog)

linkedin.com/in/kozodoi

github.com/kozodoi

n.kozodoi@icloud.com

Experience

07/2022 – present **Applied ML Scientist**, *Amazon Web Services*, Berlin.

- building generative AI solutions to address diverse business problems across industries
- leveraging AWS to design, build, and deploy custom LLM- and ML-powered solutions
- working with large-scale customers such as Nasdaq, Ryanair, and adidas [Team]

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 – 03/2022 **Ph.D. in Machine Learning, Information Systems**, *Humboldt University*, Berlin.

Published 5 papers on ML for credit risk analytics with over 230 citations. Degree with Honors. [Thesis]

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. Degree with Honors.

Selected Certifications

01/2023 **AWS Certified Solutions Architect Associate**. Showed comprehensive understanding of AWS services.

01/2022 **AWS Certified ML Specialty**. Demonstrated ability to build, train, tune and deploy ML models on AWS.

05/2020 **Udacity ML Engineer Nanodegree**. Learned to build and deploy ML models using Amazon SageMaker.

Further AWS certifications and course completions are available on my website. [Certificates]

Selected ML Projects

09/2018 – 02/2022 **Kaggle Master**. Won 18 competition medals. Reached top-1% of the global user ranking. [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]

09/2019 – 04/2021 **Fairness: Package for Fair ML**. Built and maintained an R package with over 26k downloads and 8 citations.

Implemented calculation and visualization of fairness metrics. Wrote documentation. [Blog post] [GitHub]

Further ML projects and their summaries are available on my website. [Portfolio]

Selected ML Publications

06/2020 – present **ML Blog**. Publishing AI tutorials, competitive ML tricks, and other posts. Reached 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] [Talk]

Further ML papers and talks are available on my website. [Publications]

Technical Stack

Tools  Python, Docker  VSCode, Jupyter  AWS (SageMaker, Bedrock, Lambda), CDK  Git, Asana

Libraries  Pandas, Numpy, Streamlit  PyTorch, Lightning  LangChain, Huggingface  Scikit-Learn, XGBoost

Soft Skills

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

- Key skills
- strong motivation to learn and be curious
 - attention to detail and scientific mindset
 - excellent communication skills [Public talks]
 - rich experience of working with ambiguous problems