Nikita Kozodoi

- n.kozodoi@icloud.com
- +49 176 *** 35125
- Berlin, Germany
- https://kozodoi.me (blog)
- https://linkedin.com/in/kozodoi
- https://github.com/kozodoi

Education

04/2018 – present **PhD in Information Systems**, *Humboldt University*, Berlin.
Doing research and on machine learning applications for credit risk analytics.

10/2015 – 12/2017 **M.Sc. in Economics and Management Science**, *Humboldt University*, Berlin, GPA: 1.30.
Focused on business analytics, data science and applied predictive analytics.

09/2014 – 12/2016 **M.Sc. in Economics**, *Higher School of Economics*, Moscow, GPA: 9.58 of 10 (1.21).
Focused on econometrics and time series analysis. Diploma with honors.

09/2010 – 06/2014 **B.Sc. in Economics**, *Higher School of Economics*, St. Petersburg, GPA: 9.38 of 10 (1.31).
Focused on econometrics and marketing. Diploma with honors.

09/2000 – 07/2010 **High School Diploma**, *High School 241*, Saint Petersburg, GPA: 1.00.

Work Experience

04/2018 – present **Data Science Research Associate**, *Monedo*, Hamburg.
- developing solutions to mitigate the impact of sampling bias on scoring models
- performing profit-driven feature selection with multi-objective evolutionary algorithms

07/2016 – 03/2018 **Research Assistant**, *Max Planck Institute for Human Development*, Berlin.
- collected and prepared a large number of data sets
- benchmarked managerial heuristics against machine learning algorithms
- analyzed factors affecting the predictive performance

10/2014 – 12/2015 **Research Assistant**, *Institute for Statistical Studies at Higher School of Economics*, Moscow.
- preprocessed survey data for statistical analysis
- conducted the econometric analysis of survey data

Technical Skills

Programming	**Languages:** Python, R, SQL	**IDE:** VS Code, Jupyter, RStudio
Productivity	**Deployment:** AWS Sagemaker	**Collaboration:** Git, Slack, Jira, Trello
Documents	**Office:** LaTeX, MS Office, Apple iWork	**References:** Endnote, Zotero

Packages	**Python:** numpy, pandas, sklearn, pytorch, tensorflow, xgboost, lightgbm, hyperopt, matplotlib **R:** mlr, caret, tidyverse (dplyr, ggplot2), data.table, xgboost, h2o, glmnet, plotly, knitr
Algorithms	**ML:** Boosting (XGB, LGBM, Catboost), Tree-based (RF, DT), LR, SVM, KNN **DL:** CNNs, RNNs (LSTM, GRU), BERT, Autoencoders, GANs (DCGAN, CycleGAN)
Applications	**Supervised:** classification, regression, recommendation, time series **Unsupervised:** anomaly detection, clustering, dimensionality reduction **CV:** image classification, object detection, style transfer, image generation **NLP:** sentiment analysis, text generation

Awards and Achievements

2020 Kaggle Competitions Master tier (1 gold, 4 silver and 3 bronze medals). Highest rank: 613th

2016 Prize-winner of the student research paper competition in Computer Science held by HSE

2014 – 2017 Awarded with Oxford-Russia Fund (2014/15) and E-Fellows.net (2016/17) scholarships

Certificates

06/2020 **Udacity Deep Learning Nanodegree:** building CNNs, RNNs and GANs in PyTorch

05/2020 **Udacity Machine Learning Engineer Nanodegree:** deploying ML and DL models in AWS

Software Packages

07/2020 – present **dptools**: Python package with helper functions for data preprocessing and feature engineering

09/2019 – present **fairness**: R package for computing and visualizing metrics of algorithmic fairness

Selected Competitions and Hackathons

08/2020 **SIIM-ISIC Melanoma Classification**: top-1% (5-people team, 3314 teams)

10/2019 **IEEE-CIS Fraud Detection**: top-3% (2-people team, 6381 teams)

02/2019 **Google Analytics Revenue Prediction**: top-2% (2-people team, 3611 teams)

12/2018 **PLAsTiCC Astronomical Classification**: top-5% (3-people team, 1094 teams)

08/2018 **Home Credit Default Risk**: top-4% (solo, 7198 teams)

06/2018 **Data Science Game 2018**: top-13 and special mention award (4-people team, 128 teams)

09/2017 **Data Science Game 2017**: top-9 and special jury award (4-people team, 145 teams)

Selected Publications

04/2020 N. Kozodoi, P. Katsas, S. Lessmann, L. Moreira-Matias and K. Papakonstantinou
Shallow self-learning for reject inference in credit scoring
ECML PKDD 2019 Proceedings, pp. 516-532 (https://doi.org/10.1007/978-3-030-46133-1)

09/2019 N. Kozodoi, S. Lessmann, K. Papakonstantinou, Y. Gatsoulis and B. Baesens
A multi-objective approach for profit-driven feature selection in credit scoring
Decision Support Systems, 120 (2019), pp. 106-117 (https://doi.org/10.1016/j.dss.2019.03.011)

Languages

English **Proficient user (C2)** *IELTS band 8.0 certificate*

German **Advanced user (C1)** *Language courses at HU Berlin*

Russian **Native speaker**

Skills and Interests

Key Skills
- Passionate about machine learning
- Strong motivation to learn and improve
- Excellent presentation skills
- Good at meeting deadlines

Hobbies Football, Scootering, Piano

Berlin, 06.10.2020