

# Ratmir Miftachov

Citizenship: German | +1 (929) 4857855 | [contact@miftachov.com](mailto:contact@miftachov.com) | Date of Birth: 02/17/1997

## EDUCATION

### Humboldt University of Berlin

*PhD Candidate in Statistics*

- Supervisor: Prof. Wolfgang Härdle

Berlin, Germany

*May 2021 - expected: Mar. 2026*

### Princeton University

*Visiting Researcher, ORFE Department*

- Invited by Prof. Jianqing Fan

Princeton, United States

*May 2023, Fall 2024*

### University College London

*Visiting Researcher, Department of Computer Science*

- Invited by Prof. Tomaso Aste

London, United Kingdom

*Aug. 2023*

### University of Mannheim

*Master of Science in Economics with Distinction*

Mannheim, Germany

*Sept. 2018 - Sept. 2020*

### University of Cologne

*Bachelor of Science in Economics with Distinction*

Cologne, Germany

*Sept. 2015 - Febr. 2018*

## RESEARCH PROJECTS

### Generalized Projection Flow for Regression Trees

*Work-in-Progress*

*[paper]*

- Proved theoretical guarantees for a novel data-driven early stopping rule for the regression tree. Demonstrated its ability to significantly reduce computational costs compared to state-of-the-art methods.

### Early Stopping for Random Forest Classifier

*Work-in-Progress*

- Extended previous work to the random forest classifier. Demonstrated higher accuracy for the early stopped forest than the deeply grown forest.

### Python Library for Early Stopping Methods

*Work-in-Progress*

*[Github]*

- Implemented a Python library for early stopping techniques for different algorithms, including a gradient descent variant, conjugate gradient, L2-boosting, and the regression tree.

### Shapley Curves: A Smoothing Perspective

*Published at Journal of Business and Economic Statistics*

*[paper]*

- Derived minimax rates for nonparametric Shapley curves as a variable importance measure and established a novel wild bootstrap procedure for finite sample inference.

### Risk-Premia in the Bitcoin Market

*Submitted*

*[paper]*

- Analyzed Bitcoin option data through the nonparametric pricing kernel and identified a different risk appetite compared to the S&P 500 based on a novel clustering algorithm.

## TALKS

### Statistics Lab Seminar (invited)

*Princeton University*

Princeton, United States

*Oct. 2024*

### International Conference on Computational Statistics (invited)

*University of London*

London, United Kingdom

*Aug. 2023*

### Conference Statistical Foundations of Data Science

*Princeton University*

Princeton, United States

*May 2023*

### Conference Recent Advances in Statistics and Data Science

*Rutgers University*

Rutgers NJ, United States

*May 2023*

### Conference Statistics of Machine Learning (invited)

*Charles University*

Prague, Czech Republic

*Oct. 2022*

## AWARDS

---

### **PhD Scholarship of the German Academic Scholarship Foundation (Studienstiftung)**

*Merit-based Scholarship awarded to the top 0.5% of German students*

*Since 2023*

- 4 years; approx. EUR 90,000

### **Dean's award for outstanding academic achievement**

*Awarded to the top 5% of 479 students in the Bachelor's program*

*2017*

## ACADEMIC WORK EXPERIENCE

---

### **Research Associate, Department of Mathematics**

*SFB 1294 'Nonlinear statistical inverse problems', HU Berlin*

*Aug. 2023 - Jun. 2025*

### **Supervision Master's thesis**

*Mustafa Suman on 'Early Stopping for Random Forest Classifier' (together w/ Markus Reiss)*

*2024 (ongoing)*

### **Teaching Assistant**

*'Statistical Learning and Data Science' (Master's level), HU Berlin*

*Spring, Fall 2022*

### **Associated researcher, DeSBI**

*'Deep Learning and Statistics Towards Understanding Structured Biomedical Data', HU Berlin*

*Since Mar. 2023*

### **Research Associate, Department of Economics**

*IRTG 1792 'High-dimensional nonstationary time series', HU Berlin*

*May 2021 - Feb. 2023*

### **Research Intern, German Federal Bank**

*Department General Statistics*

*Frankfurt, Germany*

*Summer 2018*

## FREQUENTLY USED SKILLS

---

**Languages:** German (native), English, Russian

**Software:** Python, R, Matlab, Git

**Mathematical statistics:** likelihood inference, parametric and nonparametric regression (e.g. splines, kernel regression), experimental design, statistical hypothesis testing, bootstrap inference (finite and asymptotic), minimax theory, probability theory, time series estimation, GMM, basics of bayesian statistics (MCMC, variational inference)

**Statistical learning:** tree based algorithms (CART, bagged trees, random forest), boosting, classification (e.g. logistic regression, k-NN, SVM), causal inference (Rubin's causal model, Pearl's structural causal model), deep learning, cluster analysis, GANs, regularization techniques (e.g. LASSO, ridge, early stopping), dimensionality reduction (PCA, UMAP, LLE, t-SNE), LSTM, explainable ML (e.g. variable importance, SHAP)