Stanislav Khrapov

Contact Information

Ossietzkystraße 6
Frankfurt am Main, Germany
Phone: +49 159 0117 9221

Email: khrapovs@gmail.com
Web: khrapovs.github.io

Personal Statement

I am a full stack Machine Learning engineer with many years of experience building data science applications starting from data ingestion, modelling, and finishing with cloud deployment and monitoring services. I am passionate about data, automation, code quality, visualisation, and communicating with both technical and non-technical stakeholders.

Professional Experience

M2hycon, Hamburg, Germany (hybrid)

Lead Data Scientist **Jan 2024 – present**

- Developed failure detection system for installation process of industrial machines.
- Replaced existing neural network model with substantially simpler, faster and more performant classification algorithms allowing for greater explainability and faster experimentation cycle.
- Refactored existing POC level code of several microservices to production level using foundational principles of software engineering.

Chintai, Frankfurt am Main, Germany

Senior Data Scientist **July 2022 – Dec 2023**

- Developed trade surveillance system based on unsupervised time series classification machine learning models. Built realistic exchange market simulation with heterogeneous traders. Wrote and open sources fast Python-based order book matching engine (OrderBookMatchingEngine). Designed a fully automated and reproducible pipeline to simulate market data, train and evaluate ML models catching illegal trading behaviour.
- Built a GitHub actions based CI/CD pipelines to release Node.js web applications interacting with the blockchain, check code quality, run unit and integration tests, search for code vulnerabilities, deploy to Kubernetes cluster running in the cloud. Organized processes across the company to streamline and speed up development and release activities starting from a PR and finishing with deployment to the production environment.
- Implemented Infrastructure as a Code management of GitHub assets using Pulumi.
- Participated in building TypeScript testing framework for Blockchain smart contacts.
- Mentored junior data science colleagues. Organized technical workshop with a purpose of active exchange of ideas and latest developments in IT.

DB Schenker, Frankfurt am Main, Germany

Data Scientist **April 2018 – June 2022**

- Designed and implemented custom time series models for forecasting of market freight prices and volumes, company internal financial indicators (EBIT, revenue, receivables, payables, etc.) to facilitate data-driven pricing decisions and liquidity planning.
- Designed and implemented recommendation system to digitize and automate ground transportation auctions in the carrier-dispatcher communication platform.
- End-to-end ML cycle including data ingestion, processing, training, forecasting, evaluation, monitoring, orchestration, and delivery software.
- Wrote software packages for automated model evaluation, comparison, and reporting.
- Produced sophisticated static (Matplotlib, Seaborn) and dynamic (Dash, Bokeh) visualizations for presentation to internal business clients.
- Worked in agile environment as a lead data scientist, scrum master, and product owner. Coordinated work of data scientists, data engineers, DevOps, business consultants. Mentored trainees and junior team members.
- Designed and performed comprehensive company-wide workshops and multi-day trainings on machine learning for audiences of up to 150 people across diverse departments, both online and offline, with topics ranging from basics up to advanced hands-on instruction.

New Economic School, Moscow, Russia

Assistant Professor of Finance **Sep 2011 – Aug 2018**

- Research: Conducted research independently as well as with co-authors in the fields of financial econometrics, option pricing, volatility modelling.
  Searched and surveyed all relevant literature on the subject. Collected, cleaned, and analyzed complex data from such sources as OptionMetrics, CRSP, TAQ, Compustat, Quandl. Collected unstructured data using web parsing methods and regular expressions. Performed Monte Carlo experiments for model selection purposes.
  Advanced knowledge of parametric and non-parametric estimation methods of non-linear models including MLE, GMM. Implemented several models and estimation methods as Python modules publicly available from GitHub repository.
  Presented at major international economics, finance, and econometrics conferences.
- Teaching: Intermediate Econometrics, Advanced Econometrics, Financial Econometrics, Data Analysis in Python. 10–50 students in each class.
- Supervision: 4–8 master and bachelor students each year.

SAS Institute, Cary, NC, USA

Summer intern **May 2007 – Aug 2010**

Programmed C module for estimation of GEE type models
Wrote examples of usage of new Copula procedure
Participated in writing of future publication "SAS/ETS User's Guide"
Added examples of usage and edited manuals for SAS/ETS procedures

University of North Carolina, Chapel Hill, NC, USA

Teaching/Research assistant **Aug 2007 – May 2011**

Teaching: Recitations in graduate classes: Intro to Probability Theory, Applied Econometrics, Advanced Econometrics.

Oregon State University, Corvallis, OR, USA

Teaching assistant **Sep 2004 – Jul 2006**

Teaching: Graduate econometrics recitations. Graded undergraduate courses in economics.

United Nations, New York, NY, USA

Summer intern **Jun 2005 – Sep 2005**

Improved annual statistical report on contributions and expenditures of member countries.

Education

University of North Carolina, Chapel Hill, NC, USA

PhD in Economics (major in Financial Econometrics) **Sep 2006 – Jun 2011**

Oregon State University, Corvallis, OR, USA

MA in Economics (major in Econometrics) **Sep 2004 – Jun 2006**

Novosibirsk State University, Novosibirsk, Russia

BA, MA in Economics **Sep 1998 – Jun 2004**

Skills

Operating systems: Linux, MacOSX, Windows
Programming languages: Python, SQL, SAS, C/C++, TypeScript
Visualization: StreamLit, Dash, Bokeh, Matplotlib, Seaborn
Orchestration: Argo, Airflow, Pachyderm
Infrastructure: Kubernetes, Helm, Pulumi, Docker

Monitoring: Prometheus/Grafana, ELK stack
ML software: Dask, SciKit, DVC
Web: Flask, FastAPI, Node.JS
Cloud platforms: AWS, Azure
Git: GitHub, GitLab
Languages: English (fluent), German (fluent), Russian (native)