Evgenii Kortukov

□ +49 176 515 78226 | @ eekortukov@gmail.com | ⊕ Website | in LinkedIn | ⍟ GitHub | ⚲ Berlin, Germany

EDUCATION

University of Tübingen   Tübingen, Germany
M.Sc. in Machine Learning; GPA: 1.1/1.0   Oct 2021 − Mar 2024
- Thesis: Studying language model behavior under realistic knowledge conflicts.

Lomonosov Moscow State University   Moscow, Russia
B.Sc. in Fundamental Informatics and Information Technology; GPA: 4.78/5.0   Sep 2016 − Jun 2020
- Diploma with Honours. Thesis: Finding similar supercomputer applications using structural pattern recognition methods.

RESEARCH EXPERIENCE

Fraunhofer Heinrich Hertz Institute   Berlin, Germany
Research Associate (PhD Candidate), XAI group   Sep 2024 − Present
- Investigating how interpretability methods can help advance the security and safety of LLM deployment.
- Supervisors: Prof. Dr. Wojciech Samek, Dr. Sebastian Lapuschkin

Tübingen AI center   Tübingen, Germany
Student Research Assistant, Scalable Trustworthy AI group   Jun 2023 − Mar 2024
- Researched LLM knowledge updating behaviors under realistic knowledge conflicts, resulting in a COLM 2024 publication. [4].
- Contributed to a qualitative study on ML practitioners' views of training data attribution methods, with findings accepted to NeurIPS 2023 Workshop XAIA [3]
- Supervisors: Prof. Dr. Seong Joon Oh, Alexander Rubinstein and Elisa Nguyen

University of Tübingen   Tübingen, Germany
Student Research Assistant, Decision Making Group   Oct 2021 − Jan 2023
- Worked on a project extending the contextual multi-armed bandit problem to the settings of distribution shifts and costly information acquisition [2]. Code is available here.
- Designed a numerical evaluation environment to model the application of multi-armed bandit algorithms to a large-scale recommender system [1]. Code is available here.
- Supervisors: Prof. Dr.-Ing. Setareh Maghsudi and Dr. Saeed Ghoorchian

WORK EXPERIENCE

Yandex   Moscow, Russia
Software Developer, Storage Engineering Department   Feb 2019 − Jun 2021
- Automated and enhanced storage cluster management system, responsible for data recovery, load balancing, and background processes that maintain fault-tolerance, robustness, and availability. This reduced time of human intervention more than tenfold, making the storage more scalable.
- Co-developed a distributed system that makes storage more scalable through managing metadata in a separate database.
- Upgraded and extended the storage statistics collection process from backend to frontend (using C++, Python, and JavaScript).

Higher School of Economics   Moscow, Russia
Intern Data Analyst, Cognitive Psychophysiology Laboratory   Aug 2018 − Jan 2019
- Developed a data preprocessing and analysis pipeline for a psychophysiological experiment (using MATLAB and Python).

PROJECTS

Tübingen AI Exhibition "Cyber and the City"   Tübingen, Germany
                                                                                         Aug 2022 − Feb 2023
- Developed and built an interactive museum exhibit showing and explaining an open-source text2image model to the Tübingen public as part of a team practical project (using Diffusers library, PostgreSQL, and a React frontend).
- The code for the exhibit is available here. The exhibition won the DFG Communicator award.

Teaching experience

Trustworthy ML   Tübingen, Germany
Teaching Assistant   *Oct 2023 – Apr 2024*
- Held tutorials and created homeworks for the Explainability part of the lecture.
- Course website

Skills

Technical: Deep Learning, NLP, RAG, Explainable AI, Classical ML, GNU/Linux, Databases, Git

Programming: Python, Pytorch, Huggingface, C++, MATLAB.

Languages: English (C2, IELTS 8.5), German (C1).

Publications

[1] Saeed Ghoorchian, **Evgenii Kortukov**, and Setareh Maghsudi. Non-stationary Linear Bandits with Dimensionality Reduction for Large-Scale Recommender Systems. *IEEE Open Journal of Signal Processing*, 2024.

[2] Saeed Ghoorchian, **Evgenii Kortukov**, and Setareh Maghsudi. Contextual Multi-Armed Bandit with Costly Feature Observation in Non-stationary Environments. *IEEE Open Journal of Signal Processing*, 2024.

[3] Elisa Nguyen, **Evgenii Kortukov**, Jean Song, and Seong Joon Oh. Exploring Practitioner Perspectives On Training Data Attribution Explanations. In *XAI in Action: Past, Present, and Future Applications*, 2023.

[4] **Evgenii Kortukov**, Alexander Rubinstein, Elisa Nguyen, and Seong Joon Oh. Studying Large Language Model Behaviors Under Context-Memory Conflicts With Real Documents. In *First Conference on Language Modeling*, 2024.