Ravil Mussabayev

+1 (206) 9391844 • Seattle, WA • ravmus@uw.edu
rmusab.github.io • GitHub • LinkedIn • Google Scholar

Education

University of Washington — Seattle, WA
Ph.D. Candidate in Mathematics, GPA: 3.81/4.0 — Sep 2018 – Jun 2024

- *Advisor*: Prof. Gunther Uhlmann
- *Research topic*: Novel Efficient and Highly Scalable Clustering Algorithms for Big Data
- Improved state of the art by more than 31.5% in accuracy and 82.4% in time for large datasets
- Published paper "How to Use K-means for Big Data Clustering" in the peer-reviewed Q1 journal "Pattern Recognition" with impact factor 8.52

University of Washington [diploma] — Seattle, WA
Master of Science in Mathematics, GPA: 3.81/4.0 — Sep 2018 – Mar 2021

Kazakh-British Technical University [diploma] — Almaty, Kazakhstan
Bachelor of Engineering in Mathematical and Computer Modeling, GPA: 3.93/4.0 — Aug 2015 – Jun 2018

- *Advisor*: Nurlan S. Dairbekov
- *Thesis*: "Self-learning neural network based traffic signal controller for an isolated intersection and construction of a new clustering algorithm in unsupervised machine learning" [thesis, presentation]

Work Experience

Research Scientist — Jun 2022 – current
Huawei Russian Research Institute — *Moscow, Russia*

- Implement, improve, and deploy code refactoring algorithms (method name prediction) for Java and C++ services used in Huawei Cloud
- Devise, train, evaluate, and deploy novel deep-learning-based models, including transformers, graph neural networks and large language models (LLMs), applied to problem of code vulnerability detection (cybersecurity) for Java, reaching unprecedented state-of-the-art final AUC score of 0.84 on real-world Java dataset
- Prepare research papers on quality vulnerability dataset collection, optimizing vulnerability detection paradigm, vulnerability representation, applying prompting and in-context learning of LLMs to vulnerability detection, as well as finetuning SOTA LLMs for in-depth vulnerability detection

Research Engineer — Mar 2015 – Dec 2017
Uniline Group LLP — *Almaty, Kazakhstan*

- Solved control problem of robotic arm in Python using input from single video camera, and published results [ieee][pdf]. Gave 2 international conference talks on project results
- Developed backend for Kazakh text-to-speech system in Microsoft Visual C++ for use in devices for people with disabilities
- Implemented Fujisaki model for pitch contour generation that resulted in more natural synthesized speech

Freelance Software Developer, *Almaty, Kazakhstan* — Sep 2010 – Sep 2014

- Developed and launched Android application Physics Lab on Google Play that reached more than 17,000 installations within 3 months
- Developed commercialized application for new cryptographic algorithm in Embarcadero Delphi
- Designed and developed desktop and Android applications for learning English words using text-to-speech technologies and MySQL database

Research and Teaching Experience

Teaching Assistant & Predoctoral Instructor — Sep 2018 – Jun 2022
University of Washington — *Seattle, WA*

- Taught introductory courses on differential equations and linear algebra for sections of 40 students
- Conducted practical college-level math classes in sections of 40 students, lead in-class discussions, held office hours, compiled and administered exams

- Tutored students in difficult math concepts, and helped them develop problem solving and analytical thinking skills

Research Assistant Mar 2018 – Aug 2018
Satbayev University *Almaty, Kazakhstan*

- Applied reinforcement learning algorithms (Q-learning, Deep Q-Network, Actor-Critic) to problem of traffic light control, and published results [ieee][pdf]
- Designed novel reward formula, which led to convergence of models in traffic simulation package SUMO, and visualized results
- Trained junior university students on basics of probability and reinforcement learning, and mentored them through research agenda

Teaching Assistant Jan 2017 – Dec 2017
Kazakh-British Technical University *Almaty, Kazakhstan*

Publications

1. Ravil Mussabayev. *Dissecting Code Vulnerabilities: Insights from C++ and Java Vulnerability Analysis with ReVeal Model*. 2023. arXiv: 2307.11454 [cs.CR].
2. Ravil Mussabayev and Rustam Mussabayev. *Optimizing K-means for Big Data: A Comparative Study*. 2023. arXiv: 2310.09819 [cs.LG].
3. Ravil Mussabayev and Rustam Mussabayev. *Strategies for Parallelizing the Big-Means Algorithm: A Comprehensive Tutorial for Effective Big Data Clustering*. 2023. arXiv: 2311.04517 [cs.DC].
4. Rustam Mussabayev, Nenad Mladenovic, Bassem Jarboui, and Ravil Mussabayev. "How to Use K-means for Big Data Clustering?" In: *Pattern Recognition* 137 (2023), p. 109269. ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2022.109269. arXiv: 2204.07485 [cs.LG]. URL: https://www.sciencedirect.com/science/article/pii/S0031320322007488.
5. Gulnur Tolebi, Nurlan S. Dairbekov, Daniyar Kurmankhojayev, and Ravil Mussabayev. "Reinforcement Learning Intersection Controller". In: *2018 14th International Conference on Electronics Computer and Computation (ICECCO)*. 2018, pp. 206–212. DOI: 10.1109/ICECCO.2018.8634692.
6. Ravil Mussabayev. "Colour-based object detection, inverse kinematics algorithms and pinhole camera model for controlling robotic arm movement system". In: *2015 Twelve International Conference on Electronics Computer and Computation (ICECCO)*. 2015, pp. 1–9. DOI: 10.1109/ICECCO.2015.7416879.

Conferences

- Talk at XII International Conference on Electronics, Computer, and Computation (ICECCO-2015) [pdf], *Almaty, Kazakhstan* Sep 2015
- Talk at XI International Asian Seminar "Problems on the optimization of complex systems", *Cholpon-Ata, Issyk-Kul, Kyrgyzstan* Aug 2015
- Reviewer and talk at XIV International Conference "Optimization and Applications" (OPTIMA-2023), *Petrovac, Montenegro* Sep 2023

Technical Skills

- **Programming:** Python (PyTorch, NumPy, Pandas, OpenCV, Scikit-learn, Matplotlib), C/C++ (OpenGL), Scala, MATLAB, R, Java, HTML/CSS, Databases (Microsoft SQL Server, MySQL), TensorFlow, Keras, Hadoop, Spark, Docker, Git, Linux shell scripts, LaTeX (wrote > 1000 pages of math)
- **Relevant coursework:** Programming Languages, Algorithms and Data Structures, Theory of Probability and Mathematical Statistics, Inroduction to Machine Learning, Introduction to Artificial Intelligence, Natural Language Processing, Deep Learning, Introduction to Cryptography, Distributed Systems, High-Dimensional Probability in Data Science, Networks and Combinatorial Optimization, Convex Analysis and Nonsmooth Optimization, Optimal Transport, Real Analysis, Topological and Smooth Manifolds
- **Engineering:** electrical engineering, computer hardware, microcontrollers (Arduino, Raspberry Pi), servo motor actuators

Leadership and Awards

- Winner (among 300 participants) of $7000 for summer research abroad by Yessenov Foundation based on leadership and research potential Jun - Aug 2017
- Held fellowship named after First President of Republic of Kazakhstan for academic excellence
- Elected member of university-wide student parliament, 10 out of 2200 students
- Winner in nominations "Best GPA 2016, 2017"
- Winner of district (2011 – 2014), city (2012, 2013), and national (2013) Olympiads in physics
- Graduated with honors from MIPT Distance Learning School of Physics and Technology (Moscow, Russia)

References

- Prof. Kenneth Bube [letter]
- Prof. Nurlan S. Dairbekov
- Prof. Gunther Uhlmann