Jafar Isbarov isbarov@vt.edu ceferisbarov.github.io U.S. (F1 visa)

Skills

Programming languages: Python, C++, Julia
Frameworks:, PyTorch, TensorFlow,, Rasa
Software development & DevOps: FastAPI, PostgreSQL, Docker, MS Azure, GCP

Education

Virginia Tech, PhD Computer Science, 2029 (expected)
The George Washington University, MS Computer Science, 2025, CGPA: 3.9 (1st of class)
ASOIU, BS Biomedical Engineering, 2023, CGPA: 3.4

Academic Experience

Graduate Research Assistant @ Virginia Tech - Blacksburg, VA Aug. 2025 –

- Working on the security of LLM-powered systems.

Visiting Researcher with Duygu Ataman @ New York University - New York City, NY Aug. 2024 – Feb. 2025

- System-level evaluation of existing prompt compression methods.
- Developing an adversarial prompt compression paradigm.
- Building TurkicMMLU - the first native multilingual language understanding benchmark for Turkic languages.

Research Assistant with Elmir Mahammadov @ Helmholtz Zentrum München - Remote Jan. 2021 – May 2021

- Reproduced two papers and performed a comparative analysis of dimensionality reduction, clustering, and visualization algorithms.
- Wrote a detailed report summarizing our findings and making recommendations for each stage of the single-cell transcriptomic data analysis pipeline.

Industry Experience

Machine learning engineer @ eiLink R&D - Baku, Azerbaijan Jan. 2025 – Aug. 2025

- Designed and led implementation of an agentic system enabling natural language queries over heterogeneous log formats in the Oil & Gas sector.

Lead machine learning engineer @ PRODATA LLC - Baku, Azerbaijan Mar. 2023 – Jul. 2024

- Built and led the company's NLP team and ML infrastructure.
- Led the development of chatbot systems in Azerbaijani.
- Designed and led the development of an internal on-premise no-code solution for chatbot development. (Stack: Rasa, Django, Nuxt, Celery, PostgreSQL, and MinIO)
- Developed an improved document retrieval strategy for hierarchical text corpora.
- Developed the SOTA monolingual foundation models for Azerbaijani.

Machine learning engineer @ Azerbaijan Artificial Intelligence Lab (within Azercosmos) Jul. 2022 – Mar. 2023

- Built a new robust spelling correction algorithm for agglutinative languages, decreasing Word Error Rate from 27.00 to 17.67.
- Built the backend system for the spelling correction solution (Stack: MariaDB, FastAPI, TensorFlow).

: Managed the data annotation team for the spelling correction project.
: Built a prototype chatbot to replace the largest customer service system (BirBank) in Azerbaijan.

Research output

: **J. Isbarov** et al. “TUMLU: A Unified and Native Language Understanding Benchmark for Turkic Languages,” https://arxiv.org/abs/2502.11020v1
: **J. Isbarov** et al., “Open foundation models for Azerbaijani language”, *ACL 2024, 1st SIGTURK Workshop,* **Honorable Mention in Best Paper category,** https://aclanthology.org/2024.sigturk-1.2/
: **J. Isbarov**, K. Huseynova “Enhanced document retrieval with topic embeddings,” *IEEE AICT 2024,* https://arxiv.org/abs/2408.10435
: **J. Isbarov** et. al., “Robust Automated Spelling Correction with Deep Ensembles”, *ACM ISMSI 2024,* https://dl.acm.org/doi/10.1145/3665065.3665070
: F. Tinner, …, **J. Isbarov** et al. “Findings of the 2nd Shared Task on Multi-lingual Multi-task Information Retrieval at MRL 2024,” https://aclanthology.org/2024.mrl-1.30/

Professional service

: **Reviewer** @ Neural Computing and Applications Journal, 1st LLMSEC Workshop (ACL 2025), 4th Multilingual Representation Learning Workshop (EMNLP 2024)
: **Officer** @ ACL Special Interest Group in Turkic Languages, https://sigturk.github.io/

Miscellaneous roles and achievements

: **Presenter** @ DevFest 2023 Baku: Machine Learning at Scale
: **Presenter** @ GDG Baku: The Challenges of the Azerbaijani Language in Artificial Intelligence
: **Co-founder** @ aLLMA Lab, where we develop open-source NLU technologies for low-resource languages
: **Open-source contributor** @ TableTransforms.jl, HuggingFace Python Client, etc.