Abror Shopulatov

📱 +998 995577753   ✉ murodbeck@proton.me   🏠 Tashkent, Uzbekistan

🔗 in/abrorshopulatov   ⭘ github.com/shopulatov   hf huggingface.co/murodbek

Profile

Highly motivated aspiring AI scientist with a strong foundation in Mathematics and a proven passion for Natural Language Processing. Leveraged five years post-high school to gain intensive hands-on experience in ML model development, large-scale data processing, and open-source contribution, focusing on advancing Uzbek language technology.

Education

High School Diploma   **Surkhandarya, Uzbekistan**
Specialized Boarding School 12   *2013-2020*

Physics and Mathematics major.

- Awarded Academic Gold Medal (GPA 5.0/5.0).
- Multiple qualifications for Republican stage Mathematics Olympiads.

Projects & Relevant Experience

Machine Learning Engineer,   **Tashkent, Uzbekistan**
(Tahrirchi)   *02/2023 - present*

- Involved in the collection and preprocessing of Uzbekistan's largest text corpus, enhancing skills in large-scale data handling and facilitating its open-source release.
- Designed and implemented an end-to-end ML pipeline for machine translation, gaining experience in parallel data collection and advanced cleaning techniques.
- Managed synthetic data generation workflows and developed a context-based translation feature, improving model performance and usability similar to Reverso.

Mathematics Tutor,   **Tashkent, Uzbekistan**
(Inspiring Education)   *02/2021 - 07/2022*

- Tutored prospective students in Mathematics and Problem Solving for elite middle school selection tests, strengthening foundational teaching skills.

Publications

- **Shopulatov, A.** (2025). *UzLiB: A Benchmark for Evaluating LLMs on Uzbek Linguistics*. Tilmoch.ai. (Blog Post). Accessed April 14, 2025, from tilmoch.ai (in Uzbek)

- Mamasaidov, M., & **Shopulatov, A.** (2024). Open Language Data Initiative: Advancing Low-Resource Machine Translation for Karakalpak. In *Proceedings of the Ninth Conference on Machine Translation (WMT24)* (pp. 606-613). Association for Computational Linguistics. (Conference Paper). DOI: 10.18653/v1/2024.wmt-1.48 / Link

- Mamasaidov, M., **Shopulatov, A.**, & Yusupov, J. (2023). *How we created an Uzbek grammar checker*. Tahrirchi. (Blog Post detailing dataset/model release). Accessed April 14, 2025, from tahrirchi.uz (in Uzbek)

Honors & Awards

- **Presidential Tech Award (1st place in AI Track)**, Republic of Uzbekistan, 2023. Awarded as part of the Tahrirchi team for contributions to Uzbek language technology and Artificial Intelligence. (Official Award Info)
- **Gold Honour**, International Youth Math Challenge (IYMC), 2021.

Skills

- **Programming:** Python
- **AI/ML Libraries & Frameworks:** PyTorch, HuggingFace Ecosystem (transformers, datasets, tokenizers), Scikit-learn, MosaicML

- **AI/ML Concepts:** Natural Language Processing (NLP), BERT, Seq2Seq Models, Machine Translation, Synthetic Data Generation, Transfer Learning, Data Preprocessing, Model Training & Evaluation, Text Corpus Development
- **Data Science Tools:** Pandas, NumPy, Matplotlib, Data Collection & Cleaning
- **Platforms & Tools:** Google VertexAI GenAI SDK, Git
- **Soft Skills:** Problem-Solving, Research, Open-Source Contribution, Project Initiative, Cross-functional Collaboration

Online Courses & Certifications

- NLP with Transformers in Python (May 2023) - Udemy
- NLP with Classification and Vector Spaces (Mar. 2023) - deeplearning.ai, Coursera
- Python for Data Science, AI & Development (Jan. 2023) - IBM, Coursera
- Data Analysis with Python (Dec. 2022) - IBM, Coursera

Languages

- **Uzbek** [Native]
- **English** [Advanced] - IELTS Overall 7.0 (L:8.0, R:7.5, S:6.0, W:6.0; March 2020)