

Akhmadjonov Mumtozbek

Github: @mumtozee

LinkedIn: Akhmadjonov Mumtozbek

Email : oltinsahifa@gmail.com

Mobile : +7 (925) 662-52-01

Telegram: @imumtozee

EDUCATION

- **Moscow Institute of Physics and Technology** Moscow, Russia
BSc in Applied Math and Computer Science Sept. 2019 – Aug. 2023
- **Moscow Institute of Physics and Technology** Moscow, Russia
MSc in Computer and Data Science Sept. 2023 – Aug. 2025

EXPERIENCE

- **Yandex Technologies** Moscow, Russia
Intern ML Developer Sep 2023 - Dec 2023
 - **ASR and AST using LLMs:** Finetuned Llama 2 in a multimodal setting using speech tokens on CommonVoice and CoVoST2 datasets. Achieved a similar BLEU score as Google's AudioPaLM-1 8B on CoVoST2.
 - **GPTQ for ASR:** Introduced SotA int4 quantization algorithm for the Seq2Seq ASR model from the production increasing it's quality (WERp -2.87% on hard samples) and inference speed and reduced the consumed GPU memory by 3 times.
- **DeepPavlov** Moscow, Russia
Junior Researcher Sept 2022 - June 2023
 - **Automatic Dialogue Graph Construction:** Developed approaches to evaluate the method on different dialogue datasets. Co-authored a paper accepted to AINL 2023.
 - **User feature extraction & inference from dialogues:** Outperformed the approach described in the article in user attribute extraction with SOTA transformer architectures.
- **DeepPavlov** Moscow, Russia
Research Intern Feb 2022 - May 2022
 - **Controllable Generation:** Achieved a similar performance with DialoGPT using special tokens compared to the approach leveraging projected attention layers to control different dialogue attributes. Reviewed dozens of articles on the related topics on ArXiv.

TECHNICAL SKILLS

- **Languages:** Python, C, C++, Java
- **Technologies:** PyTorch, PyTorch Lightning, HuggingFace Transformers, Accelerate, PEFT, trl, scikit-learn, numpy, pandas, matplotlib, Git, Docker, Flask, FastAPI, MapReduce (YTsaurus, Hadoop)
- **Knowledge:** NLP, Computer Vision, Classical ML, DL, Statistics & Probability Theory, Convex Optimization

PUBLICATIONS

- **Neuroinformatics 2022 NN workshop:** Presented a poster describing controllable DialoGPT.
- **XXV International Conference on Neuroinformatics:** Co-authored a conference paper about dialogue response selection enhancement using conversational graphs.

PROJECTS

- **RLHF for GPT2:** My attempt to align GPT2 using PPO, DPO and SFT on different settings, and serve the aligned model using FastAPI and gRPC.
- **PEFT for BERT:** My attempt to fine-tune BERT-like models on MLM task using HuggingFace PEFT.
- **Controllable DialoGPT:** Flask API to connect to a chat client that uses the model trained in this notebook.

ACHIEVEMENTS

- **Tinkoff & MSU Math Contest 2021:** winner diploma
- **Open Doors 2023 International Student's olimpiad Math & AI:** winner diploma
- **Open Doors 2023 International Student's olimpiad CS & DS:** winner diploma
- **Skolkovo Hack 2022:** Top-4 with team "ProDaters"

LANGUAGES

- **Uzbek:** Native **Russian:** Bilingual proficiency **English:** Advanced (C1) **Spanish:** Elementary (A2)

JOB INTERESTS

- **Data Science & Machine Learning Engineer/Researcher:** NLP, Computer Vision, Conversational AI, Generative models (GPT family)