Andrei Panferov

ML Researcher

☏ +7 (926) 431-76-71 | ✉ andrei@panferov.org | ⌂ BlackSamorez | in BlackSamorez

Experience

KAUST, Optimization and Machine Learning Lab   *Saudi Arabia*
Research Intern   *July 2023 - September 2023*
- Conducted research under the supervision of Prof. Peter Richtarik
- Derived theory and ran experiments on **Distributed Optimization**, focusing on **Communication Compression**

Eqvilent (HFT fund)   *Remote*
Software Engineer   *July 2022 - May 2023*

Yandex   *Moscow, Russia*
ML Engineer intern (NLP)   *March 2022 - July 2022*
- Refactored and optimized an **LLM inference framework** enabling abstract *tabular data* insertion for efficient **Map-Reduce** inference
- Increased test coverage from **0 to 85%** through rigorous unit testing
- Took part in developing a *universal LLM benchmarking solution* adapting 2 datasets for it

Terra Quantum AG   *Moscow, Russia*
Researcher   *July 2020 - July 2022*
- Designed a framework for quantum circuit optimization using
- Developed an NMR spectra analysis library for quantum computations
- Optimized Large Language Models for chat assistant applications, reducing latency by 30%

Awards

International Physics Olympiad   *Israel*
Gold Medal   *Summer 2019*

Education

Moscow Institute of Physics and Technology (MIPT)   *Moscow, Russia*
Bachelor of Science in Physics and Math   *2020 - 2024*
- Achieved a perfect 5.0/5.0 GPA
- Second minor in Teaching Methods and Pedagogy

Yandex School of Data Analysis (YSDA)   *Moscow, Russia*
Advanced Program in Machine Learning (Master's Level Courses)   *2021 - 2023*
- Specialization in DL and NLP
- Open-source contributions (see *Projects*)

Teaching Experience

Natural Language Processing, MSc level course, YSDA   *Russia*
Teaching Assistant   *Fall 2023*
- Prepared seminar homework on Model Compression
- Individually graded said homework for a class of over 300 students

Machine Learning, MSc level course, YSDA   *Russia*
Teaching Assistant   *Spring 2023*
- Graded paper reviews and assessed ML projects for course requirements

Computational Math, BSc level course, MIPT   *Russia*
Teaching Assistant   *Spring 2023*
- Prepared and delivered a 2-hour seminar on Reinforcement Learning to a class of 30 undergraduates

Open-Source Contributions

tensor_parallel

- An open-source **python library** for tensor parallel **PyTorch** models training and inference. Tight inregration with **Hugging Face transformers**
- More than *300* ⭐ on *GitHub*
- Created under the supervision of **Yandex Research** researchers

HuYaLM-100B

- An implementation of **YaLM-100B** LLM specifically for **Hugging Face transformers**
- Based on the officialy published **Megatron-LM** implementation

LLaMA implementation for transformers

- Took part in adapting the **LLaMA** model to the **Hugging Face transformers** library
- Fixed the positional embedding errors and optimized past key-value handling

NLP Bot Project

- Designed an automatic data collection system to extract thousands of dialogues from internet forums and published them as a dataset
- Refined the collected data using a pretrained sentiment analysis **BERT** model
- Fine-tuned a **GPT-2** model for **chatbot** purposes on the refined dataset
- Developed the code for a **Telegram bot** deployment, utilizing the trained model
- Published an article on **Habr** about the project, reaching the daily top-1 in the **ML** section