ALEXANDER BUYANTUEV

☎ +7 911 292 71 53 | ✉ alexbuyan.dev@gmail.com | ⌗ alexbuyan | in alexbuyan

EDUCATION

HSE University   Saint-Petersburg, Russia
Bachelor of Applied Mathematics and Computer Science   *Sep. 2020 - June 2024*
- **Relevant courses**: *Algorithms and Data Structures, C/C++, Java, Python for Web, Python for Backend, Machine Learning, Databases, Computer Networks, Calculus, Linear Algebra, Discrete Mathematics, Probability Theory, Math Statistics*
- **Yandex Data School courses**: *Machine Learning, Natural Language Processing, Reinforcement Learning*
- **AI Talent Hub courses**: *Project Management in Data Science*

WORK EXPERIENCE

RL Researcher   *Sep. 2023 - Present*
Huawei R&D, Network Scheduling Team   Saint Petersburg, Russia
Linear erasure codes construction using reinforcement learning | *Python*   Sep. 2023 - June 2024
- Developed RL-FEC software to design linear codes for different conditions and limitations of communication network
- Researched the application of different RL algorithms (DDQN, SAC, PPO) for designing linear erasure codes
- Developed **FastAPI** app to evaluate the performance of linear codes which helped to **speed up** the training up to **50 times**
- **Increased the accuracy** of evaluation from 10^{-3} to 10^{-6}
- **Compared** obtained by RL-FEC linear erasure codes with FlexFEC, LDPC and Reed-Solomon codes in BEC and Simple-Gilbert loss model
- **Showcased the advantage** of RL-FEC software as a **universal** approach for linear code construction

Software Engineer Intern   *Nov. 2022 - Sep. 2023*
Huawei R&D, Cangjie Team   Saint Petersburg, Russia
CSV support for Data-Driven Testing in Cangjie | *Cangjie*   Sep. 2023
- Implemented CsvParser in Cangjie to **parse** data from CSV files
- Developed CsvStrategy to provide data for unit tests and **contributed** it to Cangjie Test Framework

LLVM IR decompiler for Cangjie | *C++, Python, GoogleTest*   Nov. 2022 - June 2023
- Designed a tool to represent LLVM IR module in C-like format that restores packages, classes and functions from Cangjie source code to **speed up** compiler's generated code analysis
- **Created test cases** for the tool from 300+ projects on Cangjie
- Developed a parallel testing framework that runs 30 test cases with 100000 lines each under 1 minute to **fix bugs** in my tool
- **Distributed** the tool inside Cangjie Team for analysis of compiler's generated code by other developers

PROJECTS

Telegram bot with integrated RAG model ⌗ | *Python*   June 2024
- **Combined** *Llama3:8b* and *sentence-transformers* embeddings for retrieval-augmented generation
- Used *Chroma* vector database to store user's PDF files
- Developed a **Telegram bot** to upload PDF files and interact with the RAG model

Movie rating prediction model in Hadoop cluster ⌗ | *Python, Hadoop, Spark*   Dec. 2024
- **Set up** Hadoop cluster and Spark session to work with datasets
- Worked with datasets using **PySpark** to describe the data
- **Trained** SGDRegressor to predict the movie rating based on tags dataset

Multilingual Embedding-based Machine Translation ⌗ | *Python*   Sep. 2023
- Loaded **embeddings** for Russian and Ukranian
- Implemented embedding space mapping using **Linear Regression**
- **Increased** results with orthogonal transformation **SVD**
- Developed **word-based translator** from Ukranian to Russian

YouTube videos comments project ⌗ | *Python*   Apr. 2022
- **Trained GPT-2 Large** on dataset of US videos and comments
- Developed an app to interact with the model using PyTube and Gradio
- Uploaded project to **Hugging Face**

SKILLS

Programming languages: Python, C/C++, Java
Technologies and Frameworks: git, PyTorch, StableBaselines, Gymnasium, Docker, FastAPI, LangChain, SQL
Languages: Russian, English (C1)