Alexander Buyantuev

+7 911 292 71 53 | alexbuyan.dev@gmail.com | alexbuyan | 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

ML Engineer *Sep. 2023 - Present*

Huawei R&D, Network Scheduling Team Saint Petersburg, Russia

FEC codes construction using reinforcement learning | *Python, PyTorch, FastAPI* Sep. 2023 - June 2024

Duties and responsibilities:

- Developed RL-FEC software to design FEC codes for different conditions and limitations of communication network
- Researched the application of different RL algorithms (DDQN, SAC, PPO) for designing FEC codes
- Developed FastAPI app to evaluate the performance of FEC codes
- Compared obtained by RL-FEC codes with FlexFEC, LDPC and Reed-Solomon codes in BEC and Simple-Gilbert loss model

Achievements:

- **Increased the accuracy** of evaluation from $\mathbf{10^{-3}}$ to $\mathbf{10^{-6}}$
- **Speeded up** training up to **50 times**
- **Showcased the advantage** of RL-FEC software as a **universal** approach for FEC codes 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

Duties and responsibilities:

- Implemented CsvParser in Cangjie to parse data from CSV files
- Developed CsvStrategy to provide data for unit tests

Achievements:

- **Contributed** to Cangjie Test Framework

LLVM IR decompiler for Cangjie | *C++, Python, GoogleTest* Nov. 2022 - June 2023

Duties and responsibilities:

- Designed a tool to represent LLVM IR module in C-like format that restores packages, classes and functions from Cangjie source code
- 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

Achievements:

- **Distributed** the tool inside Cangjie Team for analysis of compiler's generated code by other developers
- **Speeded up** compiler's generated code analysis

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
- 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*

PDF Editor with LaTeX support | *Java* Mar. 2022 - June 2022

- Designed a converter of UI objects to PDF document to transfer project's data to PDF file
- Implemented rendering of LaTeX equations to allow users to work with math formulas
- Developed a utility to download and save user's files to enable users to save their projects
- Added font support in UI and PDF to empower the customization of documents

Messenger with Trello boards | *C++, PostgreSQL, Trello API* Jan. 2021 – May 2021

- Created PostgreSQL database to store users' information
- Implemented curl library wrapper to work with Trello API to support team boards
- Developed server's functionality to handle requests to the database

Parser generators comparison | *Python, Java, ANTLR4, Parglare* Oct. 2021

- Researched basic functionality and limitations of ANTLR4 and Parglare and compared them with other parser generators
- Compared generators' performance on ambiguous grammar recognition and collected data for the report
- Described research results in the report

Skills

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