# DAMIR ABDULAYEV

**GitHub**: ⊙ SpeedFireF

**Machine Learning Engineer**

**Languages**: Russian (Native), English (Fluent)

📞 +7 (987) 299-00-47

G abdulayevdamir2@gmail.com

## EXPERIENCE

**Junior Machine Learning Engineer** `NLP` `LLM` `BERT` `Classic ML`    June 2024 – August 2024
*MTS DIGITAL, ADP AI team*    *Russia*
- Transformed an internal service into a chatbot, using an LLM to generate and execute SQL queries, delivering cellular service data results in an easily readable chat format and via Excel export.
- Developed a lightweight, deployable BERT-like model to refine search results by filtering false positives, then aggregated and analyzed the data for accurate storage fault identification.

**Machine Learning Engineer** `LLM` `HuggingFace` `Transformers` `RecSys`    March 2024 – June 2024
*INNOPROG, ML team*    *Innopolis, Russia*
- Developed an LLM model with a recommendation system to generate personalized programming tasks, tailoring difficulty and themes to individual student learning levels.
- Created a training dataset by implementing a knowledge distillation pipeline, using a GPT model as the teacher for efficient model training.

**Data Scientist Researcher** `LLM` `HuggingFace` `Quantization` `Distillation`    May 2023 – Feb 2024
*Innopolis University, "ПИШ"LLM research team*    *Innopolis, Russia*
- Development of methods for optimizing and adapting LLMs for efficient deployment on mobile devices, web services and extensions, including the latest quantization and distillation techniques.
- Conducted research and implemented model compression techniques to optimize an LLM for generating GitHub commit messages, enabling real-time performance in a microservice with limited computing resources.

## EDUCATION

**Innopolis University**    Innopolis, Russia
*Bachelor of Data Science in Computer Science*    *Sep. 2021 – Sep. 2025*
- GPA: 4.5/5.0
- Key Courses: Machine learning, Deep learning, Natural language processing, Computer Vision, Generative Artificial intelligence, Databases, Big Data, Statistics for Data Science, Linear Algebra, Mathematical Analysis

## PROJECTS

**Games Recommender System** `RecSys` `LightFM` `Hybrid Model` `Python` `PyTorch`
- The recommendation system employs a hybrid approach, combining collaborative filtering and content-based filtering techniques to provide personalized game suggestions based on Steam Dataset.
  ⊙ Project page
  Ⓜ Medium blog post

**Music-to-image generation** `CV` `Diffusion model` `Image generation` `Sound embeddings`
- The project's goal is to convert music into visual artwork using CLIP sound embedding and Diffusion model.
  ⊙ Project page

**Hackaton "Цифровой прорыв"** `CV` `U-net` `Semantic Segmentation`
- Infrastructure object recognition using satellite data case from Scholtech. Results: 2th place (winners) overall in Beginners category
  ⊙ Project page

**Super Mario Bros AI agent** `CV` `Reinforcment learning` `Q-learning` `open-ai gym`
- This project focuses on training an AI agent to play the popular game Super Mario Bros using reinforcement learning techniques.
  ⊙ Project page

## TECHNICAL SKILLS

**Data Science**: Python, Pandas, NumPy, Matplotlib, Seaborn, SciPy
**Machine Learning**: Scikit-learn, NetworkX, Catboost, LightGBM, XGBoost, opencv
**Deep Learning**: PyTorch, Huggingface, NLP, Transformers, CV, RecSys
**MLOps**: PostgreSQL, neo4j, MongoDB, PySpark, Docker, GIT