

ABYLAY DALABAY

Almaty, Kazakhstan

+7-701-730-19-46 dalabayabilay@gmail.com [telegram](#) [GitHub](#)

Education

Kazakh-British Technical University

Bachelor of Engineering and Technology in Computing Systems and Software

Aug 2021–Jun 2025

Almaty, Kazakhstan

Yandex School of Data Analysis

Additional Education in Machine Learning Engineering

Sept 2024–Jun 2026

Remote

Selected among 2500+ candidates (250 chosen for Remote Study) after rigorous testing, exams, and interviews.

First semester completed courses:

- * **Algorithms and Data Structures:** Most interesting problems: Tree isomorphism, FKS hashing, SQRT decomposition, AVL, Fenwick trees
- * C++ solutions of contests: [Algo repository](#).
- * **Machine Learning:** Implemented from scratch: linear regression, logistic regression, decision trees, convolutional layers, batch normalization, backpropagation
- * **C++ Programming:** Implemented from scratch: vector, deque, shared and weak pointers, functional language interpretator. [C++ repository](#)

Experience

Junior NLP Engineer

Bank Center Credit

May 2024–Present

Almaty, Kazakhstan

Multilabel Classifier for Customer Complaints

- * Processed 25K dialogues, removing operator templates.
- * Used DBSCAN to filter 7K non-complaint cases.
- * Auto-labeled minor classes via ChatGPT-API.
- * Fine-tuned **Multilingual-e5-large**, achieving $F1 = 0.8$ across 80 classes.

Reranker for the Chatbot

- * Built a reranker with Multilingual-e5 on 6K examples.
- * Dataset contains 5 hard positives and negatives for each answer
- * Improved Recall@10 by 20% over cosine similarity.

Kaz-Rus Language Detector

- * Augmented 150K Kazakh comments to 600K transliterated pairs.
- * Trained **FastText** (84% acc) and **LogReg** on **BERT** embeddings (98% acc).

Projects

YSDA Kaggle Dota2 Competition | *sklearn*

Feature engineering with logistic regression

Ranked **Top-80/250** participants, achieving ROC-AUC **0.74**.

[Kaggle repository](#).

NLP | *PyTorch*, *HuggingFace*

Implemented RNN, LSTM, embeddings-CNN, different sampling techniques, and a custom tokenizer.

[NLP repository](#)

Technical Skills

Languages: Python, SQL, C++

Libraries: PyTorch, OpenCV, Numpy, sklearn

Achievements

Expert (1795 rating) on [Codeforces](#)

Top-17 in Northern Eurasia Finals (ICPC Semifinals) 2023 (Honourable mention)

Nov 2023

Bronze Medal in International Zhautykov Olympiad in mathematics

Feb 2021

Silver Medal in Asian-Pacific Ocean Olympiad in mathematics

Mar 2021