Danil Kizeev

Keywords

python, optimization, machine learning, data science, computer vision, NLP, statistics, numpy, pandas, PostgreSQL, ClickHouse, scipy, transformers, pytorch, C++, Rust, Docker.

Experience

2021 Apr – 2021 Jun **Data Science intern @ Hyprr**, Saint-Petersburg
Solved case of automatic categorization of users posts in social network. Scraped, labeled the data. On the first iteration used Google API for getting images descriptions. On the second iteration used CLIP over ResNet and DistilBERT for the posts, KMeans for clusterization and mapping objects.

2022 Sep – 2023 Mar **Game developer @ Hive**, Perm, Russia
Applying genetic and differential evolution algorithms to build an efficient matchmaker for multiplayer in online mobile games; making visualizations via egui and macroquad; building data collection pipelines; rust language. Making player and debug cameras; building raycasting.

2023 Oct – 2025 Feb **AI Engineer @ ECMC Exponenta**, Moscow, Russia
Working with the problem of searching competitors to the main product. Made our own labeling framework using telegram bot API. Trained CLIP for russian language using CC12M+COCO+translations datasets+own translations using Helsinki-NLP models. Used CLIP over ResNet and DistilBERT. Using the resulting labels trained siamese model and aligned it to classification problem. Made models with FashionCLIP and XGBoost as a head for classification, compared them. Implementing optimal prices models using contextual multi-armed bandits (CMAB). Creating models of trend, seasonality, demand using time-series models (Prophet, XGB, xARIMAx, linear) and neural networks (BiLSTM, Transformers). Applying to production with customer. Lifting using Docker+Flask for MVP's

Education

University

2017 – 2018 **Cybersecurity department, Bachelor**, *Tver State University, TVGU, Russia*

2018 – 2021 **Applied mathematics and computer science, Bachelor**, *Saint-Petersburg State University, SPBGU, Russia*

2023 – 2025 **Data Analytics, Master**, *ITMO University, Russia*

MOOCs, CSC, YSDA

Apr 2018 - Jun 2018 **Algorithms: theory and practice. Methods**, https://stepik.org/course/217

Jun 2020 - Jun 2021 **Machine Learning Course**, https://mlcourse.ai

2022 **Mathematical statistics @ CSC,**

Sep 2022 **NLP course @ YSDA,**

Sep 2022 **Deep CV & Graphics course @ YSDA**

Saint Petersburg, Russia
+79040283459 • *kizeevdanil@yandex.ru*

Projects

Study projects/pet projects

Mindmaps for everything
https://github.com/breadfan/mindmaps-for-everything
Smart diagrams for math-related objects
Created mindmaps for better understanding ("Zettelkasten" inspiration) and keeping in mind tech-related objects. There are statistics, NLP, optimization and deep learning using pytorch now. Maps are available in **rus** and **eng** languages.

Lyric music recommender
https://github.com/breadfan/lyrics-based-songs-recommender
Using doc2vec and BERT embeddings for recommendation based on songs lyrics.
Created models with doc2vec and DistilBERT, created UMAP reduced embeddings for 2-d and 3-d visualisations (Plotly). Generated bot using «telegram» library.

Automatic posts categorization
https://github.com/breadfan/Bachelor-Thesis
Applying BERT for automatic posts categorization in social network.
Using BERT, BERTopic and word2vec + TF-IDF for labeled images/video categorization.
Upgraded quality of models from BERT to word2vec.
Having categories need to make mapping from labels amount of posts to that categories for making simple recommendations.

Accelerated MDM-method
https://github.com/breadfan/Accelerated_MDM_method
Researching acceleration of an MDM-method
As a course work for third year two methods were implemented: MDM and accelerated MDM methods with visualization for 2- and 3-dim. Cases for running time comparison.

Russian-Karelian translator
https://github.com/breadfan/kielven_project
Creating translator for livvi based on NLLB, mBART and Qwen3. Creating framework for parallel corpora labelling
As a masters degree created first online karelian translator. Elevated Meta's results, as well as Alibaba (for Qwen). Parsed corpora from Russian-Karelian dictionary. Parsed olo.wikipedia.org. Parsed vepkar.ru. Made frameword for corpora parallel labelling.

Technical Skills

Languages Python, Rust, plpgsql, C/C++
VCS Git
OS Windows, Linux (Ubuntu)

On the Internet

GitHub https://github.com/breadfan
StackOverflow https://stackoverflow.com/users/9850300/taciturno
LinkedIn https://www.linkedin.com/in/rocauc
LeetCode https://leetcode.com/breadfan/
Telegram https://t.me/rocauc