Petrov Oleg

oepetrov0@gmail.com | github/necroshine0

EDUCATION

BSc. Applied Mathematics and Informatics, Machine Learning Specialization Moscow | Enrolled 2019
Higher School Of Economics, Faculty Of Computer Science
Teacher Assistant (TA):
Adaptive math for data analysis, Sep – Nov 2021
Python for data collection and analysis, Feb – Mar 2022
Industrial development tools, Apr - June 2022
Machine Learning, Sep – Jan 2023
Coursework:
Simulating the mouse behavior using discrete operators on a graph;
Educational analytics: curriculum analysis

PROJECTS

CURRICULUM CLUSTERING NLP, clustering, text processing, labeling , inner metrics
I have collected a dataset of documents annotations of educational courses of various universities and clustered it. Used TF-IDF, Word2Vec and Doc2Vec models to get documents' embeddings, implemented inner Xie-Beni metric for performance estimation. Labeled the dataset and implemented B-Cubed metric as a scoring.

DATA SCIENCE, KAGGLE data science, machine learning, analytics
I am into solving data science problems at Kaggle and getting practical experience in analytics and machine learning. This repository presents my hard skills and work results.

ACTIVITY ANALYSIS python, data collection, EDA, Visualization
I was collecting and processing data to classify the type of human movement: standing, walking, running, going upstairs, riding a bicycle, car, scooter, ect. Performed EDA, processed outliers, used FFT to decompose accelerator's tracks into frequencies. Got the 5th place from 84 on the leader board with $accuracy = 0.71232$ using primitive if-else model.

DATABASE BOT python, telegram, SQL, database
I developed the telegram bot performing definite queries to the implemented database (SQLite3). Collaborated in the design of the conceptual database model, DDL and DML.

DISCRETE OPERATORS ON A GRAPH C++, graph algorithms, oop, neurobiology
The project under the supervision of the Laboratory of Comparative Physiology of Higher Nervous Activity of Animals, Moscow State University. The goal was to modeling mouse's behavior in a maze using discrete operators (DOs). I designed DOs' algorithms using inheritance and polymorphism, implemented Floyd-Warshall algorithm, developed an method constructing a sequence based on the probability distribution of DOs, implemented the probabilities' selection method on a grid by the Levenshtein distance. The goal was achieved but the model was not very stable.

TIC-TAC-TOE ONLINE java, REST-service, oop-design, casual game
I designed product vision, user stories and UML-class model. Have been developing a client application (frontend): HTML-requests and interface.

SKILLS

Hard:
Programming: Python, C++
Data workflow: numpy, pandas, SQL, nltk
Visualization: matplotlib, seaborn, folium (geo)
ML: scikit-learn, catboost/xgboost, statsmodels
DL: pytorch, gensim; a little experience with transformers (BERT)
Soft:
Good at team work, communicative, stress-resistant, adaptive, pedantic

HOBBIES

Sports, guitar, history, photography