Petrov Oleg

oepetrov@edu.hse.ru | 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
Coursework: Simulating the mouse behavior using discrete operators on a graph;
Educational analytics: curriculum analysis

STUDY 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 to get documents' embeddings, implemented inner Xie-Beni metric for performance estimation. The next step is dataset labeling and performance estimation via Dunn's Index or B-Cubed metric.

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.

OTHER python, machine/deep learning, statistics
This repository contains my homeworks related to machine learning, also including linear algebra, basic of matrix computations and numerical methods, statistics, optimization methods. There are also some Kaggle competitions, e.g. New York City Taxi Trip Duration, Weather in Szeged 2006-2016, etc.

SKILLS

Hard:
Languages: C++, Python, Java (basics), C (basics), SQL
Machine Learning: Linear models, decision tree, bagging, gradient boosting, EM-algorithm, kernel models, clustering
Deep Learning: Perceptron, convolution models, pre-trained BERT transformer, word2vec, doc2vec, etc.
Math: Linear algebra, calculus, matrix computations, discrete math, theory of probability and statistics, optimization methods
Soft:
Good at team work, communicative, stress-resistant, adaptive, pedantic