Kuleshova Polina

Student

Higher School of Economics (HSE)
+7-926-011-19-51
pokuleshova@edu.hse.ru
polindromka@gmail.com

Responsible and inquisitive 3rd-year student of the Faculty of Computer Science of the Higher School of Economics. With experience in applying machine learning on real biological data for generating new data samples and/or clustering data. Enjoys researching new information and applying new methods to solve problems.

Education

Higher School of Economics (HSE) / Faculty of Computer Science (School of Software Engineering)

(2019-2023)

- **Language.** Upper-intermediate level
- **Estimates.** Top 6% in the rating (average score 9.5 (“excellent”), https://www.hse.ru/en/ba/se/ratings?course=3&from=555130949)
- **Course work.** Course work for last year was with the Sberbank AI Lab on automatic interpretation of an ECG by Holter. We were faced with the task of clustering QRS complexes, the initial data had R-peak markings and were preprocessed. we developed an autoencoder that was supposed to increase clustering performance. It was marked as one of the best projects in the direction and will be published on the website in this section (https://cs.hse.ru/en/cppr/best_projects_eng). A good clustering was obtained, but it was not possible to achieve high homogeneity (data from different clusters were assigned to one, based on the markup of doctors).
- **Minor.** As part of a minor (additional track), I am going through bioinformatics - the analysis of genomic sequencing data. We are taught to use and analyze databases to obtain new information, independently search and study the task. And we use statistical methods for analyzing big data.
  Tech stack: Biopython
- **Additional courses.** Computer vision, Mathematical methods for data analysis, Analysis of sequencing data, CS231n course (CS231n: Convolutional Neural Networks for Visual Recognition, Stanford)

Experience

Sberbank AI Lab (AI in medicine)/ Trainee

(October 2021 - nowadays)

- Improving the classification of the ECG signal using generative adversarial neural networks (GANs). Now our goal is to generate 8 seconds of a 12-lead ECG and add them to the main

database to improve the performance of the classifier. We generated a pretty good signal and we are going to write an article.

Tech stack: Python (PyTorch, NumPy, Pandas, SciPy, Scikit-learn, Matplotlib, Plotly, Seaborn), Linux, GIT

International Laboratory of Algebraic Topology and Its Applications / Summer internship
(July-August 2021)

- Determination of sleep stages by EEG signal. We build attractors, persistent diagrams and find Betti numbers and apply the Mapper algorithm to them. After analyzing the results obtained, several pure clusters were identified, which made it possible to assume that this method is effective.

Tech stack: Python (Giotto-TDA, MNE, NumPy, Pandas, Matplotlib, Plotly, Scikit-learn), GIT