Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Bogdan Kirillov

PhD Program: Life Sciences

Title of Thesis: Uncertainty Quantification and Neural Network Interpretation for studying CRISPR mechanics

Supervisor: Assistant Professor Maxim Panov

Name of the Reviewer:

I confirm the absence of any conflict of interest	Date: 13-09-2023

Reviewer's Report

Adaptation of actual mathematical methods for different biological problems is one of the most important areas of interest in the life sciences, as it allows us to discover new patterns among the organisms being studied. The main idea of the work presented for review was the application and implementation of machine learning, as well as the new methodology invention for the errors estimation of the results obtained during the machine training. The issue of analyzing method errors is critically important because a significant difficulty in using machine learning methods in biology and medicine is the frequent receipt of obviously incorrect results, following which can lead to unexpected consequences, especially if they are used in clinical medicine.

The basis of the research is two independent experiments, the common part of which is the mathematical apparatus developed by the author and adapted to the tasks. One of the experiments was the task of classifying the structure of skin in melanoma, and the second was a direct search and classification of Cas9 and Cas12a events that can occur during DNA editing in several thousand different organisms. When reading the thesis, one gets the feeling that the experiment on classifying the structure of the skin in melanoma is not directly related to CRISPR-Cas mechanics, but the author deliberately includes this subsection in the text of the work as a justification for the performance of the mathematical methods used.

In the work there are used the most advanced machine learning methods, such as Capsule Neural Networks, Deep Neural Networks, Explainable Machine Learning and others. All analyzed data sets were freely available and well characterized. In fact, a purely analytical and computational study was carried out. However, in the text of the work there is no information of the hardware characteristics of the computer equipment used. In such works, they usually form a separate chapter in the "materials and methods" section, where all software and hardware characteristics are described, at least briefly. If there is no description of the characteristics, it can be very hard to reproduce the work published.

Another nuance is that when comparing data obtained by different consortia, there is not only directly measured error, but also summarized error in conclusions, which includes the type of protocol used to

generate the final data loaded into open databases. Of course, when solving a specific task, such indirect errors can be neglected, but most likely they would be worth mentioning in the text.

These comments are technical and do not affect the quality and practical value of the presented work. In addition, the scientific significance of the results obtained is beyond doubt. The developed method can be successfully applied in any other industries where there is a high risk of negative learning, and then, as a consequence, obtaining erroneous results.

The dissertation submitted for review is a good and fully completed scientific work. The structure of the thesis is classic, containing all the necessary sections. The literature review takes up no more than 30% of the total work. The materials and conclusions of the work were published in highly rated peer-reviewed journals and fully include all the approaches described in the work. The work was done by the author independently, which allows the assignment of the expected academic degree to the author.

Provisional Recommendation

☒ *I recommend that the candidate should defend the thesis by means of a formal thesis defense*

☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report*

☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*