Exploring Machine Learning Interpretability by Analyzing Tumor Suppressor Genetic Sequence Data

Todorov, Hristo (School: High School of Mathematics and Natural Sciences "Professor Emanuil Ivanov")

With the application of machine learning techniques to various fields (for example, computer vision and healthcare), the problem of interpretability is gaining importance. Building transparent models is critical in the context of computational biology as they could be used to identify underlying biases and fairness issues as well as to extract novel biological insights through understandable model representations. We created machine learning approaches for analyzing raw tumor suppressor genetic sequence data while focusing specifically on determining reference genes from randomly extracted k-mers, which is a challenging task due to the data sparsity. Our results suggest that the encoding of the input data has a strong impact on the representations the models learn and that SHAP values are a useful tool for interpreting the behavior of convolutional neural networks trained on limited genomics data.

Awards Won: