Predicting Lung Cancer Survival and Treatment Using Machine Learning

Rao, Kevin (School: Moscow State School No. 1537)

The objective of this investigation is to determine the most accurate machine learning algorithm for recommending surgical and radiation treatment and predicting survival for lung cancer. Machine learning innovations and national cancer registries have made cancer-applied machine learning much easier. Because of this and the fact that medical errors are the third most common cause of death, artificial intelligence trained with machine learning has the potential to make fewer errors than humans and have the medical experience of many doctors and other artificial intelligences. A hypothesis was formulated stating that if the decision tree classifier was used to recommend surgical and radiation treatment, and if the multilayer perceptron classifier was used to predict survival, the two algorithms would be the most accurate. A program written in python and scikit-learn was used to determine the accuracy of the Gaussian naïve Bayes, support vector classifier, decision tree classifier, and multilayer perceptron classifier for recommending radiation and surgical treatment and predicting survivability. The results revealed that the SVC had a greater accuracy than all other algorithms in all experiments. A t-test was performed on the data, revealing that the results were significant for all comparison except for GNB vs. MLPC in the survivability experiment. The research hypothesis was not supported by the results. This is likely due to the SVC using more resources and creating a lenient decision boundary. This research may contribute to future studies that investigate the application of deep learning to cancer and the analysis of special biological markers.