Dr hab. Vitaliy Yakovyna, Uniwersytet Warmińsko-Mazurski w Olsztynie

Software Defect Prediction using ML («Przewidywanie defektów oprogramowania przy użyciu uczenia maszynowego»)

Due to the rapid increase in software development, the coding quality issue is becoming more and more critical. Since the timely detection of errors in software is a manual, resource-consuming, and not always reliable process, scientific interest in its automation is growing, and, in particular, the issue of defect prediction using machine learning methods is relevant. This work reviewed existing approaches to software defect prediction and proposed a new classification model based on a combination of existing approaches. The developed model is based on stacking, where each base classifier undergoes a learning process using individual data balancing and feature selection methods. The model was tested on data from the PROMISE repository. Classification metrics such as accuracy and F-scores were used as a check. The application of the developed model made it possible to achieve a classification accuracy of 0.839 and F-score - 0.909, which is better than the average result of "simple" classifiers. As a result, research on the dataset from the PROMISE repository have shown that using the proposed approach allows for improving the classification metrics during the prediction of software defects.