

## Abstract

Empirical studies on software defect prediction models have come up with various predictors. In this study we examined variable regularized factors in conjunction with Logistic regression. Our work was built on eight public NASA datasets commonly used in this field. We used one of the datasets for our learning classification out of which we selected the regularization factor with the best predictor model; we then used the same regularization factor to classify the other seven datasets. Our proposed algorithm Variant Variable Regularized Logistic Regression (V V RLR) and modified V VRLR; were then used in the following metrics to measure the effectiveness of our predictor model: accuracy, precision, recall and F - Measure for each dataset. We measured above metrics using three Weka models, namely: Bayesian Logistic Regression, Naive Bayes and Simple Logistic and then compared these results with V V RLR. V RLR and modified VVRLR outperformed the weka algorithms per our metric measurements. The V V RLR produced the best accuracy of 100.00%, and an average accuracy of 91.65 %; we had an individual highest precision of 100.00%, highest individual recall of 100.00% and F - measure of 100.00% as the overall best with an average value of 76.41% was recorded by V V RLR for some datasets used in our experiments. Our proposed modified V V RLR and variant V VRLR algorithms for F - measures outperformed the three weka algorithms.