Abstract

The aim of this research is to predict fault in the process of producing important Android applications using data mining techniques. Predictable models must not only be correct in predicting fault, but also must be understandable, which needs the user to understand the motivation of the model prediction. Unfortunately, understandability of the fault types is ignored in order to achieve the predictable efficiency of the fault prediction models. In order to solve this problem, some trees are extracted from the random forests and support vector machines for the logistic regression and the rule extraction algorithm are used; also, NASA MDP data are used for extracting the model. The method of evaluating the prediction of software fault is the use of ALPA algorithm and extraction of random trees (RF) and (Logistic) regression for weka software. In the method of creating trees, (REPTree and C4.5) and black box model (Logistic, RF) are used. The results show that the trees extracted from the black box models discuss the prediction way of black box which is more understandable as well as more correct than the trees from direct work on the data.
References

## Index Terms

| Computer Science | Information Sciences |

## Keywords

Fault prediction, software effort, Android, data mining