Abstract

According to data from the ministry of communication and informatics, in 2018 active smartphone users in Indonesia more than 100 million people and in the same year data from the android user statcounter in Indonesia as much as 90.85%. The high use of android makes the android operating system a target for malware attacks. Malware is a system that is programmed to infiltrate an operating system, an operating system that has been attacked by malware can be damaged and even with more malicious intentions malware can be used to steal important data. The vulnerability of malware attacks and can harm android users so further analysis is needed, because the existing problems encourage this research to be done with the aim of early detection. In this case, a machine learning approach is used to classify android malware attack data. The machine learning used in this study is support vector machine (SVM) and random forest. Both machine learning methods were chosen because in previous studies it was proven that both methods are very effective at classifying with high accuracy. In this paper, a comparison between the Support Vector Machine (SVM) method and the Random Forest method in classifying data, as well as comparing the results of accuracy with previous research. The classification process using the Support Vector Machine (SVM) method produces a precision value of 97%, a recall value of 97%, and an f1-score value of 97%, and an accuracy of 96.23%, in the Random Forest method it produces a precision value of 99%, a recall value of 99%, an f1-score value of 99%, and an accuracy of 98.99%. According to the results of the experiment, the Random Forest method is superior to the Support Vector Machine (SVM) method and the approach proposed in this study has a higher matric performance result above 95% than previous studies.

Keywords : Android, Malware, Machine learning, Support Vector Machine (SVM), Random Forest.