Abstract
Code clones are portions of code that are identical or similar to others. Detecting code clones can help find redundancy in code. The problem with detecting code clones manually is that the reported code clone results are not always consistent. One way to increase the F1-score in the process of detecting code clones is to automate the process of detecting code clones. Detecting code clones automatically using the token approach method represents source codes as a sequence of tokens which allows detecting code clones with different line structures. The token approach method is a heavyweight method because it requires a special parser language for each different programming language. This study aims to create an automatic code clones detection system using the token approach method to analyze and compare the F1-score value in detecting code clones. This research was conducted on three application source codes, namely Apache, Postgre, and Python with the F1-score results being "0.94", "0.55", "0.89". The results of this study obtained the highest F1-Score compared to three other code clones detection tools, including Simcad, Nicad, and MeCC. The implementation of the code clones detection system in this study shows that the system can detect code clones better than the other three methods.

Keywords: code clones, automatic code clones detection system, token approach method.