ABSTRACT

Person's personality can be predicted from Twitter data by building a model using machine learning. The learning algorithm used can be one algorithm or combination of several algorithms such as Ensemble. Stacking is one example of Ensemble. Stacking has a structure consisting of two level learners making it possible to use several different classifiers to improve the performance of system. However, Stacking has weakness. It performs worse in multi-class dataset than two class dataset. Previous study has shown that adding level learners in the original Stacking structure can improve the performance of system in either multi-class or two class dataset. This study proposes a method called Modified Stacking, which adds one level learner in the original Stacking structure so that it becomes three level learners and adds grammatical features to improve the performance of Personality prediction system. The evaluation results using 10-Fold cross validation indicate that the system obtained average prediction accuracy of 99.62% for Modified Stacking and 99.24% for Stacking. It can be concluded that Modified Stacking has better performance in terms of accuracy than Stacking. In addition, the additions of level learners and grammatical features have an effect to improve accuracy. However, the execution time increased significantly over the addition of level learner in the Stacking structure.

Keywords: Ensemble Learning, Stacking, Modified Stacking, Neural Networks, Grammatical Features