Abstract

Twitter is one of the most popular social media to interact with users. In addition, Twitter is used to get the latest information about a product, socio-political issues, celebrities, and so on. Twitter users can repost the information by retweeting a tweet so that the information can be spread again to other users. This study aims to build a retweet prediction system and see how the performance and accuracy of the Naïve Bayes classification method are based on two feature models, content-based and user-based. Dataset splitting uses k-fold cross validation with a value of k = 10. The results obtained in this study are good enough with an accuracy rate of 76.41%, precision rate of 76.72%, recall rate of 99.41% , and f1-score of 86.42%.