Abstract

Twitter is one of the most popular social media platforms for spreading information in real time. Users love this platform for its Tweet and Retweet feature. With Tweets, users can compose messages and share them directly with their followers. Messages can contain text, media such as images, videos, or URLs. Meanwhile, Retweet allows users to re-share other users' tweets with their followers. Retweet feature is considered an effective way of spreading information as evidenced by the high number of retweets, which shows that the information in the tweet is spread quickly and widely. The aim of this study was to predict the number of retweets based on several features, namely user-based, content-based, and time-based. The classification method used in this study is an artificial neural network, which is optimized by the Cuckoo search algorithm. The evaluation results show that the developed model has an accuracy of 75, an accuracy of 76%, a recall of 87% and a f1-score of 81 without managing class imbalance. Under undersampling conditions, the model's accuracy dropped to 69%, with 72% accuracy, 63% recall and 67% f1-score. Under oversampling conditions, the accuracy reached 70%, with an accuracy of 68%, recovery of 75% and a f1-score of 72%. These results show that the use of user, content and time-based features, as well as the application of the Cuckoo search algorithm-optimized artificial neural network classification method, is effective. In predicting retweets.
Keywords: ***Retweet, Twitter, Cuckoo Search, ANN, Undersampling, Oversampling*****.**