The Influence of Sentiment on the Movement of Bank Mandiri (BMRI) Stock Price with Word2Vec Feature Expansion and the Naïve Bayes-Support Vector Machine (NBSVM) Classifier

Ridhwan Nashir[1], Erwin B. Setiawan[2], Didit Adytia[3]
Informatic School of Computing, Telkom University
Indonesia
[1]ridhwanashir@student.telkomuniversity.ac.id
[2#]erwinbudisetiawan@telkomuniversity.ac.id
[3]adytia@telkomuniversity.ac.id

Abstract

***Abstract*—Sentiment towards a company is suspected of influencing the company's stock price movement. The sentiment is gathered from Twitter, Youtube, Facebook and news media CNBC, Kontan, Detik, CNN, Stockbit, and Liputan6 discussing Bank Mandiri. Word2Vec is used to reduce vocabulary errors in sentiment analysis using word embedding. The Word2Vec model was built using the combined corpus of Wikipedia articles and scraped data with a total of 474,277 lines of text data. This study indicates that the correlation between sentiment and stock movements of Bank Mandiri has a positive correlation with a low relationship, indicated by the Spearman Rank test coefficient value of 0.138 and 0.123 for positive and negative sentiment, respectively. The NBSVM classification model outperforms the NB and SVM methods, where the baseline NBSVM gets an accuracy of 64.67%, and after the feature expansion process, the accuracy becomes 70.42%, an increase of 5.75%. This study proves there is a correlation between sentiment and the movement of Bank Mandiri's shares, and Word2Vec feature expansion can increase the model's accuracy.**

Keywords—data mining, sentiment analysis, nbsvm, naïve-bayes, support vector machine, word2vec, rank spearman, stock prices