Analisis Sentimen Pengguna Kereta Cepat di Sosial Media Menggunakan Word2Vec dan Model Peluang Naïve Bayes yang Terbaik

Muhammad Dinan Islamanda[1], Yuliant Sibaroni[2]

[1,2]Fakultas Informatika, Universitas Telkom, Bandung
[1]dinanislamanda@student.telkomuniversity.ac.id, [2]yuliant@telkomuniversity.ac.id

Abstract
By using the Twitter microblogging feature, users can post short tweets with limited characters that express their thoughts and opinions regarding a matter. The newest transportation in Indonesia, a high-speed train namely Whoosh is one of the things that Twitter users responded to. This latest transportation has led to the emergence of opinions from the Indonesian people which are shared publicly in various media, one of which is social media. Therefore, to make it easier for business people or companies to understand public opinion regarding service improvements in the future, sentiment analysis on social media is needed to determine user opinions regarding high-speed train transportation. In this research, sentiment analysis of high-speed train users will be carried out on social media Twitter using Word2Vec and Naïve Bayes as classification methods. In this research, a comparison of Naïve Bayes models will also be carried out to find out the best Naïve Bayes method opportunity model. Simultaneously, the Word2vec feature extraction method was chosen because Word2Vec can be used to improve model performance and increase the accuracy of sentiment classification. This research found that the Word2Vec Skip-Gram model outperformed the Word2Vec CBOW model. The best model obtained was the use of the Gaussian Naïve Bayes and Word2Vec Skip-Gram models with an accuracy score of 77.18%, precision 70.35%, recall 76.09%, and f1-score 73.10%.

Keywords: high-speed train, naïve bayes, sentiment analysis, twitter, word2vec