

Feature selection based on particle swarm optimization algorithm for sentiment analysis classification

Vivine Nurcahyawati^a, Zuriani Mustaffa^b

^a Universitas Dinamika, Faculty Technology And Informatics, Department Of Information System, Surabaya, Indonesia

^b Universiti Malaysia Pahang, Faculty Computing, Department Of Computer Science, Pahang, Malaysia

ABSTRACT

Online media serve as a potential secondary data source for studies on sentiment analysis. The current conditions of the data sources are very different, and it offers a variety of writing systems. Therefore, the results of accuracy in sentiment analysis are very important. An improved approach was proposed to increase the sentiment analysis accuracy based on text pre-processing and Naïve Bayes Classifier algorithm hybrid with Particle Swarm Optimization (NBC-PSO). Furthermore, the proposed algorithm solves the complex background problems about noise data and feature selection that affect the classification performance on sentiment analysis. This proceeded with the classification of positive or negative sentiments on these texts using NBC. Subsequently, the feature selection based on PSO was created to improve the accuracy. The experimental results showed that the proposed approach has a significant effect on sentiment score and polarity detection.

KEYWORDS

English language; Indonesian language; Naïve Bayes Classifier; Particle Swarm Optimization; Sentiment analysis; Text pre-processing

REFERENCES

1. C. R. AYDIN and T. GÜNGÖR, "Combination of Recursive and Recurrent Neural Networks for Aspect-Based Sentiment Analysis Using Inter-Aspect Relations", *IEEE Access*, vol. 8, 2020.
2. I. P and E. D. Windasari, "Sentiment Analysis on Travel Destination in Indonesia", *Proc. of 2017 4th Int. Conf. on Information Tech. Computer and Electrical Engineering (ICITACEE)*, 2017.

3. V. Nurcahyawati and Z. Mustaffa, "Online Media as a Price Monitor: Text Analysis using Text Extraction Technique and Jaro-Winkler Similarity Algorithm" in *Emerging Technology in Computing Communication and Electronics (ETCCE)*, Dhaka, 2020.
4. L. Widiastuti, "Pemilihan Fitur Pada Analisis Sentimen Review Travel Online Menggunakan Algoritma Naïve Bayes Dalam Penerapan Mutual Information Dan Particle Swarm Optimization (PSO)", *Indonesian Journal on Computer and Information Technology*, vol. 3, no. 1, 2018.
5. M. Kartiko and Sfenrianto, "Accuracy for Sentiment Analysis of Twitter Students on ELearning in Indonesia using Naive Bayes Algorithm Based on Particle Swarm Optimization", *Journal of Physics Conference Series*, vol. 1179, 2019.