PROJECT REPORT

OPTIMIZING SENTIMENT ANALYSIS OF FOOD AND DRINK REVIEWS ON TOKOPEDIA USING SVM ALGORITHM: A COMPARISON OF GRID SEARCH AND RANDOMIZED SEARCH METHODS

TEGAR JAYA YUSTIAN MAHA DEWA
21.K1.0048

Faculty of Computer Science
Soegijapranata Catholic University
2024

ABSTRACT

Reviews written by shoppers have valuable information. In this day and age, many people shop online through marketplaces. These reviews have sentiments or emotions that can be used as a source of information to improve the standard of services or products provided by the seller. Hence, sentiment analysis becomes important. To perform sentiment analysis, within a review there is an algorithm of choice. Many other studies have used the Support Vector Machine algorithm to perform sentiment analysis. However, the SVM algorithm is highly dependent on the parameters that are selected. Therefore, optimizing these parameters is very important. In other research, techniques such as Grid Search and Random Search are quite often used. Therefore, in this study, researchers want to compare the two optimization techniques to find out which parameters are more accurate. This research compares which optimization parameter has higher Accuracy, Recall, Precision, and F1-score results using a dataset of 4061 reviews on Tokopedia obtained from the Kaggle.com site. The results obtained in this study are: Grid Search produces an accuracy value of 0.7562, slightly higher than Randomized Search, which has an average accuracy of 0.7531. In terms of precision, Grid Search gives a result of 0.7281, while Randomized Search has an average precision of 0.7258. In terms of Recall, Grid Search is again slightly superior with a value of 0.7562, while Randomized Search has an average recall of 0.7531. In terms of F1-score, Grid Search has an F1 Score of 0.7419, while Randomized Search has an average F1 Score of 0.7393. For computation time, Grid Search takes about 502.71 seconds, while Randomized Search is much more efficient with an average time of 129.06 seconds.

Keyword: Sentiment_Analysis, Support_Vector_Machine(SVM), Parameter_Optimization, Grid_Search, Randomized_Search.