Abstract

The Kanjuruhan disaster on 1 October 2022, gained the peoples attention. People share their thoughts on social media. Their posts contain a variety of perspectives. Sentiment analysis is possible to use on a dataset of people's posts. This final project applies the supervised learning Support Vector Machine (SVM) method with feature expansion using Word2Vec and TF-IDF as weighting. Three SVM kernels—rbf, linear, and polynomial—are applied. Three split data techniques and two different types of training data are used to train each kernel. Training data with oversampling and training data without oversampling are the two types of training data. The best result gained from using rbf kernel, split ratio 70:30, and oversampling. From it, oversampling trained model have relatively stable in every split rasio and kernel without having significant difference.

Keywords: sentiment analysis, kanjuruhan disaster, football, twitter, machine learning