Abstract

Social media is a popular network through which users can share their reviews about various topics, news, products, etc. People use the internet to access or update reviews, so it is necessary to express opinions. Sentiment analysis is to classify these reviews based on their opinion as either positive or negative categories. First, we preprocess the dataset to convert unstructured reviews into a structured form. Then we use a lexicon-based approach to convert structured reviews into numerical score values. In the lexicon-based approach, we preprocess the dataset using feature selection and semantic analysis. Stop word removal, stemming, POS tagging, and calculating sentiment scores with the help of SentiWordNet dictionary have been done in preprocessing. Then, we apply the classification algorithm to classify opinions as either positive or negative. Support vector machine algorithms are used to classify reviews where the RBF kernel SVM is modified by its hyperparameters which are the soft margin constant C and Gamma γ. Optimized SVM gives better results than SVM and naïve Bayes. At last, we compare the performance of all classifiers with respect to accuracy.
References


Index Terms

Computer Science    Information Sciences

Keywords

Sentiment analysis, Text mining, SentiWordNet, SVM, Naïve Bayes, RBF kernel SVM