An optimized multi-layer ensemble framework for sentiment analysis

ABSTRACT

Public opinion plays an important role in decision making tasks of various fields. Sentiment Analysis is a key task in summarizing sentiment opinions as it classifies opinion documents according to its sentiment group of positive and negative. Machine learning based classification is efficient and versatile. The ensemble concept is used to improve classification accuracy by combining the decision of multiple classifiers. In this work, a framework for sentiment analysis is designed to extend the concept of ensemble upon all subtasks of machine learning classification in order to achieve better analysis. There are 3 subtasks in machine learning based sentiment analysis which are feature extraction, feature selection and classification. The ensemble concept is applied to all 3 tasks by combining different methods to perform the tasks and combine their results. optimization is performed by using Genetic Algorithm to find the combination of methods that could perform better. The proposed framework is tested on 4 different domain datasets and the sentiment analysis accuracy is shown to be very high. Future works includes testing the framework on different domains of classification and different optimization algorithm.