Abstract

Sentiment analysis is a recent advance in text mining applications for analyzing textual data according to orientation of human comments to determine whether they are positive, negative, or neutral. Different data mining techniques and algorithms such as support vector machine, naïve Bayes, decision tree, k-nearest neighbor and other techniques are used for analyzing textual data. These techniques are evaluated based on Arabic language due to its richness and diversity that can lead to difficulties in analyzing and mining large number of morphological and linguistic words that can lead to different meaning. This research provides sophisticated categorization of most or all recent articles according to the algorithms used in analyzing sentiment data. A comparison table for the proposed algorithms is presented that explains each algorithm and its use in mining and analysis of Arabic textual data and provides different evaluation for each sentiment analysis and classification algorithm according to different categories such as sentiment type, feature selection, sentiment polarity, domain oriented, data scope and data source, algorithm used in the sentiment analysis or classification, and the best algorithm result during the analysis and mining process. The experimental results explain that
support vector machine algorithm presents high accuracy with approximately 77% when compared to other text mining algorithms. Different algorithms of sentiment analysis and classifications are evaluated based on their use in Arabic language which has not been evaluated before.

References


Index Terms

Computer Science  Algorithms

Keywords

Sentiment Analysis, Sentiment Classification, Arabic Textual Data, Text Mining, Support Vector Machine