

Sentiment Analysis and Opinion Mining for Business Products on Twitter using Deep and Machine Learning Techniques

By

Mosaab Emad Al-Sheikh Daher

Supervisor

Dr. Tarek Kanan

Al-Zaytoonah University of Jordan, 2022

Abstract

In recent years, social networking sites have exploded in popularity. People utilize social media platforms to express their opinions on virtually all topics. These opinions are expressed in a variety of formats, including blogs, tweets, Facebook postings, online message boards, Instagram posts, etc. Sentiment analysis involves computationally characterizing and categorizing the opinions expressed in comments, blog posts, and tweets. Typically, the objective of sentiment analysis is to determine a customer's opinion on a product or service. Customer feedback is essential for organizations, and social media, a potent platform, may be leveraged to improve and expand company potential provided timely analysis of social media feedback is performed. Consequently, we are concentrating on analyzing the customer reviews of numerous restaurants around the Arab world. To categorize each comment as positive, negative, or other, we employed preprocessing tools for the Arabic language and three well-known NLP techniques (Normalization, Stop Word Removal, and Stemming).

We also compiled datasets containing around 1,804 publications from Arabic social media platforms Twitter and Facebook. In ML, we employed a well-known multiclass classifier; the most well-known are Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (J48), Random Forest (RF), and K-Nearest Neighbor (KNN) (KNN). We determined the accuracy of the classifier by computing F1-Measure, Recall, and Precision. In contrast, we employed Deep Learning

techniques, which include Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN). Word Embedding is a component of the feature extraction approach that has been implemented (Word2Vector).

The results showed that the SVM classifier is given the highest value of F1-Measure for both Twitter datasets with (82.7%) and the Facebook dataset with (81.5%). Using deep learning techniques, we get results better than ML classifier as DNN gives the highest value than the other neural network techniques in the Twitter dataset, we get the value of F1-measure with (86.8%) and in Facebook dataset with (85.8%).

Keywords: Sentiment Analysis, Machine Learning, Natural Language Processing (NLP), Restaurant Reviews Classification, Arabic Social Media.