World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:14, No:07, 2020

StockTwits Sentiment Analysis on Stock Price Prediction

Authors : Min Chen, Rubi Gupta

Abstract : Understanding and predicting stock market movements is a challenging problem. It is believed stock markets are partially driven by public sentiments, which leads to numerous research efforts to predict stock market trend using public sentiments expressed on social media such as Twitter but with limited success. Recently a microblogging website StockTwits is becoming increasingly popular for users to share their discussions and sentiments about stocks and financial market. In this project, we analyze the text content of StockTwits tweets and extract financial sentiment using text featurization and machine learning algorithms. StockTwits tweets are first pre-processed using techniques including stopword removal, special character removal, and case normalization to remove noise. Features are extracted from these preprocessed tweets through text featurization process using bags of words, N-gram models, TF-IDF (term frequency-inverse document frequency), and latent semantic analysis. Machine learning models are then trained to classify the tweets' sentiment as positive (bullish) or negative (bearish). The correlation between the aggregated daily sentiment and daily stock price movement is then investigated using Pearson’s correlation coefficient. Finally, the sentiment information is applied together with time series stock data to predict stock price movement. The experiments on five companies (Apple, Amazon, General Electric, Microsoft, and Target) in a duration of nine months demonstrate the effectiveness of our study in improving the prediction accuracy.

Keywords : machine learning, sentiment analysis, stock price prediction, tweet processing
Conference Title : ICMCMA 2020 : International Conference on Mobile Communications and Multimedia Applications
Conference Location : Paris, France
Conference Dates : July 20-21, 2020