

Abstract

In this thesis, we look beyond extracting binary sentiment in regards to News Headlines and Tweets. As a data source, we target tweets and headlines from well-known financial newspapers, explicitly addressing the top 5 Big Tech companies. To examine the effectiveness of sentiment and Ekman's emotions in predicting future stock price movements, we develop multiclass emotion and sentiment classifiers utilizing a supervised learning approach. Moreover, we manually annotate our corpora for positive, negative, and neutral sentiment as well as one of Ekman's emotions: anger, joy, surprise, sadness. We did not confirm any robust correlation between daily stock price movements and the distribution of sentiment and emotions. However, we did observe that tweets are less neutral than news headlines. Finally, we implement a simple investing strategy by extracting sentiment polarity scores using VADER and other metrics such as followers and shares. Two classifiers, SVM and ANN, delivered robust predictions for Google and Amazon compared to weak predictions for the rest of the companies. Nevertheless, the results suggest that sentiment polarity can effectively predict future stock price movements compared to finer-grained emotion classification.

JEL Classification C53, G41, G17, C61

Keywords News Headlines, Tweets, Sentiment Analysis, Emotions

Title Stock Market Prediction: A Multiclass Classification on Emotions and Sentiment Analysis for Tweets and News Headlines