

Yaxin Li

Feb 11, 2020

Summary of *A Robust Predictive Model for Stock Price Prediction Using Deep Learning and Natural Language Processing* by Sidra Mehtab and Jaydip Sen.

In this paper, it presented several approaches to stock price and movement prediction on a weekly forecast horizon using eight regression and eight classification methods. These models are based on machine learning and deep learning approaches.

Data: Two authors built, fine-tuned, and tested these models using daily historical data of NIFTY 50 from January 2, 2015 till June 28, 2019. The raw data is suitably pre-processed and suitable variables are identified for building predictive models.

Algorithm Training: After designing and testing the machine learning and deep learning-based models, the predictive framework is further augmented by bringing in public sentiment in the social media in addition to the historical stock prices, as the two inputs a fuzzy neural network-based SOFNN algorithm.

Evaluation: The **classification** models are tested using the following metrics: (i) sensitivity, (ii) specificity, (iii) positive predictive value, (iv) negative predictive value, and (v) classification accuracy. For the **regression** methods, we use mean Absolute Percentage Error (MAPE), and the correlation coefficient between the actual and predicted values of the response variable (e.g., close_norm) as the two metrics.

Conclusion: The study has conclusively proved that public sentiments in social media serve as a very significant input in predictive model building for stock price movement.

Summary of *STOCK TREND PREDICTION USING NEWS SENTIMENT ANALYSIS* by Kalyani Joshi.

This project is about taking non-quantifiable data such as financial news articles about a company and predicting its future stock trend with news sentiment classification.

Data and Algorithms: Assuming that news articles have an impact on the stock market, this is an attempt to study the relationship between news and stock trend. Three different classification models were chosen for data training and prediction. They were Random Forest (RF), SVM and Naïve Bayes.

Evaluation: This project evaluated all three classifiers performance by checking each one's accuracy, precision, recall, ROC curve area. The conclusion shows that RF and SVM perform well in all types of testing while Naïve Bayes is not as good as the other two. Experiments are also conducted to evaluate various aspects of the proposed model and encouraging results are obtained in all of the experiments. The accuracy of the prediction model is more than 80% and in comparison with news random labeling with 50% accuracy.

Conclusion: Random Forest worked very well for all test cases ranging from 88% to 92% accuracy. Accuracy followed by SVM is also considerable around 86%. Naive Bayes algorithm performance is around 83%. Given any news article, it would be possible for the model to arrive on a polarity which would further predict the stock trend.