International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2017): 7.296

Stock Price Movement Prediction using Attention-Based Neural Network Framework

Kartik Goyal[1], Nitin Bansal[2], Soumyabrata Kundu[3], Ayan Kundu[4], Nitish Jain[5]

[1]IIT Delhi, Department of Engineering Physics, Hauz Khas, New Delhi 110016, India

[2]IIT Roorkee, Department of Earth Sciences, Roorkee - Haridwar Highway, Roorkee, Uttarakhand 247667, India

[3]ISI Bangalore, Stat Maths Unit, 8th Mile, Mysore Road, RVCE Post, Bengaluru, Karnataka 560059, India

[4]ISI Bangalore, SQC & OR Unit, 8th Mile, Mysore Road, RVCE Post, Bengaluru, Karnataka 560059, India

[5]Financial Services, Innoplexus Consulting Services Pvt. Ltd.,
7th Floor, Midas Tower, Rajiv Gandhi Infotech Park, Hinjawadi, Pune, Maharashtra 411057, India

Abstract: ***There is a lot of scientific work going on NLP trying to predict the impact of news on a stock price, much of this uses basic features (such as bags-of-words, named entities etc.), but fails to capture structured entity-relation, and hence lacks accuracy. 1) Encoding the information like daily events, meta-stock information and stock's 50 days moving average using LSTM. 2) Employing attention mechanism to rate the relevancy of all events for each stock. 3) Using non-linear neural network on the weighted events to predict the stock movement. The model achieved an accuracy of around 72% on test set.***

Keywords: Stock Price Movement, Neural Network, Attention Mechanism, NLP

1. Introduction

The stock price movement prediction is the act of trying to determine the behaviour of a stock in the near or distant future given some data and occasions relating to the stock. The well-known efficient-market hypothesis (EMH) suggests that stock prices reflect all currently available information and any price changes based on the newly revealed relevant information.

Stock market prediction is the area of extreme importance in the field of finance.In finance, stock market and its patterns are extremely un- stable in nature. Investors and market analysts study the relation of market news with the trend in the movement of market behaviour and used to plan for buying or selling their invests accordingly .News articles can play an important role in predicting the movement of the stock prices as movement of the market prices depends on the response of the investors to the news articles.

Previously,there is lot of scientific work trying to predict the impact of news on a stock price, much of this uses basic features (such as bags- of-words, named entities etc.), and doesn't capture structured entity- relation information, and hence lacks accuracy.

We have built an attention based neural network model (ATT-NN) to predict the impact of a news on a stock using news articles, stocks meta information and stocks moving average as inputs. Then the inputs are encoded in their respective representational formats using long short term memory cell and computing the correlations between news articles and stock. We then predict whether a stocks price will increase or decrease by 0.4% of the current value or will have impact less than the specified percentage.

2. Problem Statement

We treat our problem as a nominal classification problem where we predict the impact of all the events happening in a single day on a particular stock as up , down or unaltered. Consider a stock, given are the events E, meta-information (includes company's about us, sector, industry and products) and current stock price p with stock prices of previous 49 days. Now the stock on the following day is q, if q - p is greater than 0.004 times p, we predict up, if lesser than then down else unaltered.

3. Methodology

3.1 Data Collection

We collected the news articles from an archive openly sourced by Reuters.com. Stocks meta information was obtained from a product of Innoplexus Consulting Services Pvt. Ltd. and the stock prices were downloaded from Yahoo Finance, which gives the historical opening and closing prices, Since this was the best grained information effortlessly accessible to us, we utilized an 50-day window in endeavoring to relate news articles to stock costs. We predict the stock movement of the companies listed at the German Stock Exchange. News articles are then pre-processed (removal of stopwords, numbers, punctuation, etc.).

3.2 Encoding all the inputs

Further, all the inputs are then encoded in a 100 dimensional vector using Recurrent Neural Network with long short term memory (LSTM). So we get a matrix of numbers of size number of events in a single day X 100 (E ={e1, e2, ..., en} to {x1, x2, ..., xn} where xi is a 100 dimensional vector), a 100 dimensional vector containing meta-information (c) and

a 100 dimensional vector containing stocks moving average with a look back of 49 days.

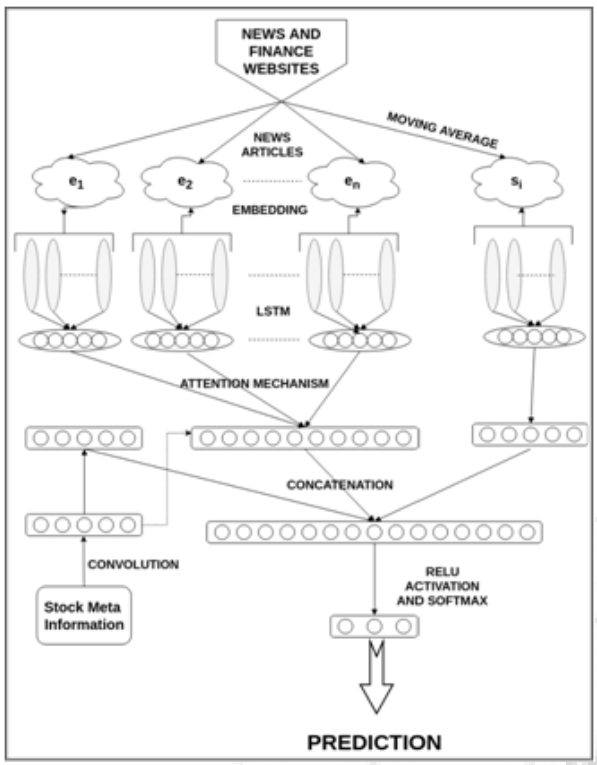

Figure 1: Attention Based Neural Network Framework

3.3 Attention based Event Relevancy Calculation

Encoded events are then concatenated and passed to a attention based neural network with two hidden layers with 300 and 100 neurons respectively having linear activation and finally to a soft- max layer which outputs a relevancy of all news articles on a stock s on a scale from 0 to 1.
Suppose after encoding the news are represented as x1, x2, x3, ..., xn where n being the number of news articles in a day, stocks meta information as c1,c2,...,cm where m being the number of stocks for which we are predicting and stock prices moving average as s1, s2, ..., sm.

$$P_{i,j} = W_r[c_i ; x_j] + b_r \quad (1)$$

$$R_{i,j} = Softmax(P_{i,j}) \quad (2)$$

$$A_i = R_{i,j} \times x_j \quad (3)$$

3.4 Prediction

Now, we concatenate Ai , stocks meta-information and stock prices moving average and pass it to softmax layer with 2 hidden layers having non-linear activations and predict 3 possible outcomes :- up , down , no impact .

$$H_{1_i} = Relu(W_r[A_i ; c_i ; s_i] + b_r) \quad (4)$$

$$H_{2_i} = Relu(W_r[H_{1_i}] + b_r) \quad (5)$$

$$O_i = Softmax(W_r[H_{2_i}] + b_r) \quad (6)$$

3.5 Model Training

We use the cross-entropy loss as the loss function for model optimization. The overall loss can be represented as:

$$J(\theta) = -\sum_i \sum_j \log(O_{i,j}) + \lambda\theta \quad (7)$$

Where represents all the parameters of our model, Oi,j is the corresponding probability output by the model for stock i in a specific date j. λ is the regularization weight. We use Adam as the optimization method, which gives the best result among all optimization methods in this dataset.

3.6 Implementation and Hyper-Parameter Selection

As you know deep learning models depend highly on the values of the hyper parameters in the model. We used a 5-cross validation to select the best fitted hyper parameter.
- Hidden Neurons in LSTM Layer = 100
- Linear Hidden Dimensions = 100
- Non-Linear Hidden Dimensions = 150
- Batch Size = 10
- λ Regularization Parameter = 0.001

3.7 Learning Curve

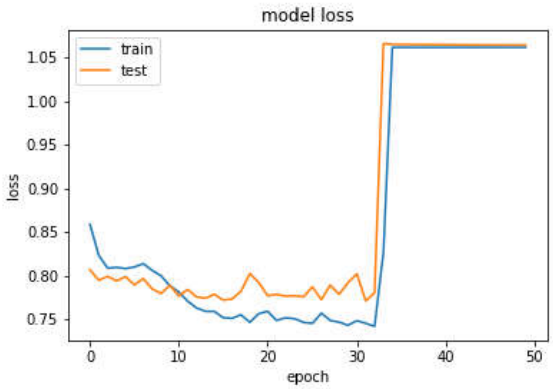

Figure 2: Learning Curve (loss vs epochs)

3.8 Baseline Models

- **RandomGus** is a basic baseline model that ignores all the information totally. It simply outputs -1 (down) or 0 (no impact) or 1 (up) at a chance of 33.33%. We also use this model to check the balance of the dataset.
- **Support Vector Machine(SVM)** is another baseline model we used. We implemented SVM model based on TF-IDF on the news text and achieved an accuracy of around 62%.

4. Conclusion

In this paper, we propose ATT-NN model to exploit the implicit cor- relations between world events and the movement of stock prices. We conduct extensive experiments on a real-world dataset, and experimental results show the superiority of our model. Therefore, we reject our null hypothesis with a accuracy of around 72% on test data.

References

[1] Liu J., Chen Y., Liu K., Zhao J. (2017) Attention-Based Event Relevance Model for Stock Price Movement Prediction. In: Li J., Zhou M., Qi G., Lao N., Ruan T., Du J. (eds) Knowledge Graph and Semantic Computing. Language, Knowledge, and Intelligence. CCKS 2017. Communications in Computer and Information Science, vol 784. Springer, Singapore

[2] Singh, R. & Srivastava, S. Multimed Tools Appl (2017) 76: 18569. https://doi.org/10.1007/s11042-016-4159-7

[3] M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*, Kuala Lumpur, 2016, pp. 322-327.

[4] K. Khare, O. Darekar, P. Gupta and V. Z. Attar, "Short term stock price prediction using deep learning," *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, 2017, pp. 482-486.

[5] Z. Hu, J. Zhu and K. Tse, "Stocks market prediction using Support Vector Machine," *2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering*, Xi'an, 2013, pp. 115-118.

Author Profile

Kartik Goyal is a third year undergraduate student in IIT Delhi. He is pursuing his majors in Engineering Physics. He has worked at Innoplexus in Financial Services Department as a Data Scientist in the past.

Nitin Bansal is a third year undergraduate student at IIT Roorkee. He is pursuing his majors in Geological Technology. He has worked at Innoplexus in Financial Services Department as a Data Scientist in the past.

Soumyabrata Kundu is a 1st year student undergraduate in Indian Statistical Institute Bangalore. He is pursuing a bachelor course in mathematics. He has worked at Innoplexus in Financial Services as a Data Scientist in the past.

Ayan Kundu is a second-year postgraduate student of Indian Statistical Institute.He is pursuing Master of Science in Quality Management Science.He has completed his graduation in Mathematics from St.Xavier's College, Kolkata.He has worked at Innoplexus in Financial Services Analytics as a Data Scientist in the past.

Nitish Jain is the VIce-President of Financial Services at Innoplexus Consulting Pvt Ltd. Former VP at Hitachi.