A Modified Deep Learning Enthused Adversarial Network Model to Predict Financial Fluctuations in Stock Market

Jasmine Sabeena, P.Venkata Subba Reddy

Abstract: **Predicting financial fluctuations in the real-time stock market is considered to be a major problem due to dynamic changes in financial data. With the advent of using artificial intelligent techniques in the context of predicting the patterns, artificial neural networks have drawn the attention of various researchers to implement the same in several computational applications. Addressing this problem, a modified adversarial network based framework is proposed with the integration of gated recurrent unit and convolution neural network. The main objective of this model is to acquire data from online financial sites and to process the obtained information using adversarial network to generate predictions.**

Keywords : *Stock Market Prediction, Generative adversarial networks, Neural Networks, Artificial Intelligence*

I. INTRODUCTION

Automated stock market analysis and prediction of the financial fluctuations in the stock market is remained as a challenging scenario for many researchers due to its dynamic nature. Even with the existence of several solutions based on the machine learning algorithms still here is a plenty of research gap in terms of accuracy in the prediction. Algorithms based on the artificial intelligence that includes pattern recognition, Natural language processing and Machine learning plays a vital role in the context of prediction in any kind of applications. Stock markets are generally influenced by various aspects that includes Sports, National and international politics, Business growth of the peer industries, climatic conditions etc. The key aspects to predict the stock market includes, Extract and analyze the share market values that are dynamically updated every minute, Develop ontology based semantic learning model using machine learning algorithms and train the system, develop an accurate prediction algorithms using deep leaning that enhances the computational efficiency and prediction performance.

  Jasmine Sabeena *, Research Scholar, Dept. of CSE, S.V.U College of Engineering., S.V. University, Tirupati, India, jasmine539@gmail.com .
  Dr. Poli Venkata Subba Reddy , [2]Professor, Dept. of CSE, S.V.U College of Engineering., S.V. University, Tirupati, India, pvsreddy@hotmail.co.in

GloVe model[1] utilized to identify the relation among words like equivalent words, items relation, communication division and cities, etc by distance between words. Predicting the stock market is the most basic part in assignment of finding the right time to purchase and sell the share. This paper propose the technique which specific GloVe vector for using the word tokenization and Convolution neural networks (CNN)[2] for efficient stock prediction that scores sentiment analysis from public domain like real time news. Finds out the correlation of positive sentiment movement and original stock value curve for accuracy with given time [3]. Coefficient between multi-variants improved by using canonical correlation analysis which provides the relationship among variables [4].

II. LITERATURE SURVEY

The main objective of this section is to analyze various studies that have addressed the problem of analyzing the fluctuations and dynamic changes in the stock market and generating future predictions based on the formulated analysis. Several research studies have addressed the automated predictions of the stocks markets and derived several machine learning based but these studies suffer in the context of accuracy due to the inconsistency of the dataset while devising a learning model [5]. Rosdyana Mangir Irawan Kusuma [6] developed a decision support framework based on the deep neural network based solution along with candlestick charts for the traders to predict and suggest the future fluctuations in the stock price that protects the investors shares. In the context of devising this model, the author have utilized a variant of neural network algorithms that includes the residual neural network, convolution neural network along with the visual geometry group network. Initially the historical data is derived from YAHOO finance application using this data they developed candlestick chart that serves as a feed input to the proposed deep neural network model. Based on the implementation evidences the researchers claimed that the proposed framework outperforms several state-of-art mechanisms when compared with the existing work. Bidirectional LSTM is integrated with CNN to devise an enhanced deep learning model that concatenates multiple pipeline of CNN in [7] to generate predictions. The purpose of using Bidirectional LSTM is to analyze the temporal data whereas CNN is the best in the context of feature extraction.

Retrieval Number F9011088619/2019©BEIESP
DOI: 10.35940/ijeat.F9011.088619
Journal Website: www.ijeat.org

2996

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

A Modified Deep Learning Enthused Adversarial Network Model to Predict Financial Fluctuations in Stock Market

The researchers claim that when CNN is integrated with Bidirectional LSTM it outperforms the traditional SVM regression in terms of accuracy. The main limitation of this approach is that sentiment analysis is not taken into consideration for predicting stock price and trends. Mustafa A.Sarawala [8] proposed a technique called SEC Reparative (textual data) that explains the computing power to invest in stock market using Data Science. Author demonstrated how a deep learning model trained on text in SEC document filings could provide a valuable signal in decision making.

This article demonstrate the surface of how the latest natural language processing techniques and deep learning models could be used to extract meaningful information from SEC reparative (textual data) and asses savings in a company's stack price. This article model can be improved by adding sentiment analysis and company news as features to see their effects on improving the model. Mohammad Obaidur Rahman [9] in the context of addressing the problem of predicting the prices of stock market developed GRU based neural network in which it follows the strategy of gated units with Boolean integers 0 and 1. In his research studies by adopting GRUs he had mitigated the local minima by changing the internal structure of GRUs. This influences the reduction of computational complexity while integrated with gradient descent algorithm .The internal structure of GRU is modified with the help of update gate and reset gate in which the process of selecting the required input data is accomplished. presented in the paper evaluated based on the parameters that effect the computational complexity in terms of learning the factors evaluation of the effect, based on mini batch gated descent and the effect of training period. Based on simulations the problem of local minima within the GRU neural networks has been exponentially reduced in his studies. Koby Todors [10] discussed about the linear canonical correlation analysis for financial data prediction through the framework called measure transformed canonical correlation analysis(MTCCA) which accomplish the task of performing the joint probability distribution on specification of MT-functions . Author explained about the MTCCA with the comparison to LCCA for nonlinear system with features having long term associations with the company trades. In the research study addressing the problem of Stock predicting Using Convolution Neural Network [11]. Authors have proposed a method that includes the deep learning model with Convolution Neural Networks to extract the features of stock prices for Chinese stock market. In this one-dimensional convolution (conv1d) function in the Convolution layer. The stock data pre-processed and given as input to process in CNN model to make predictions in 1D sequential and shown the result as reliable.

III. PROPOSED MODEL

The main objective of the proposed work is to design and develop a hybrid model which is embedded with the process of collecting data from online news resource such as yahoo finance, Indian finance news and MSN money etc. Further the collected data is tokenized based on the sentiment analysis, the output from the sentiment analysis is processed using extended RCNN [12] such that it includes two layers in which comparison and prediction of the stock market values will be processed. The final output layer will provide exact prediction set of dynamically extracted stock market.
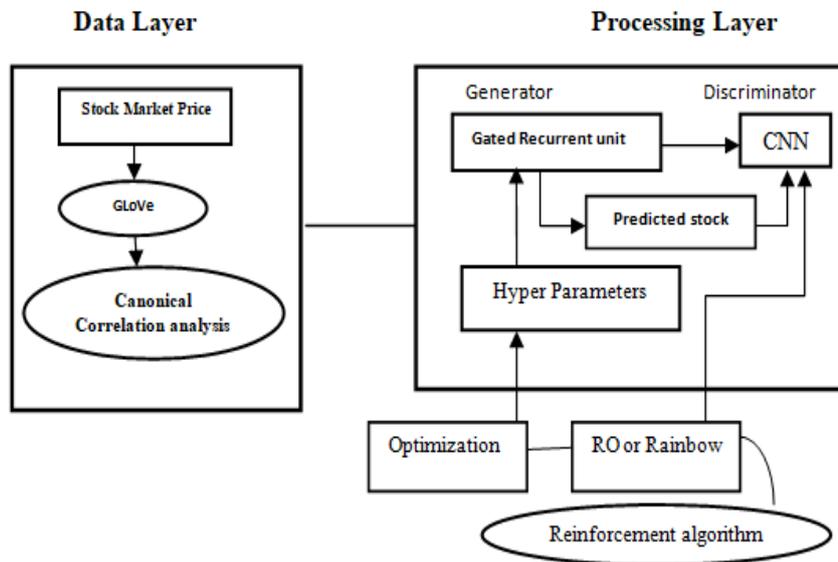

Figure 1. Proposed Model

A. Data Extraction from online news databases:

In this context the main objective is to fetch the dynamic data from the server in which the sense and nifty data values are extracted using python programming language. Initially, this data set are processed to perform sentiment analysis such that to identify the pattern and fluctuations of the Sensex data and derive predictions from the historical news data.

The process of extracting online news data includes the following steps

- Initially the required packages for python programming are pre-installed in Ubuntu 16.04 LTS server.

- In this scenario we will be using python lxml library in the context of processing online XML or HTML web data related to news databases.

- Parsing the online news data bases will derive the resultant set of the required raw data on which the sentiment analysis could be initiated.

B. Tokenizing data for sentiment analysis:

Glove [1] is the statistical model which processes based on the unsupervised learning algorithm in the context of obtaining the vector representation from words. The working idea behind this framework is to evaluate the ratio of the probability with relative to word occurrences that are considered from corpus.

In our model GloVe model [1] is adopted as technique to analysis the word occurrence in the stock market data gathered from online news database. The word to word occurrence in the GloVe model [1] is demonstrated using following instance. For the convenience of the demonstration the following sentence is considered from yahoo finance data [13], "inflows of news business showed the smallest rise seen this side of the global financial crisis ". The ratio of probability and word occurrence in this sentence is analysed through GloVe model by considering the covariance based probability model as derived below:

Table 1 Probabilistic model to analyse word occurrences using GloVe model

	Low	Rise	Constant
P(A/Global Finance)	$2.0 * 10^{-2}$	$1.7 * 10^{-2}$	$1.1 * 10^{-3}$
P(A/Business news)	$1.8 * 10^{-3}$	$3.0 * 10^{-4}$	$1.9 * 10^{-6}$
P(A/Global Finance)/P(A/Business news)	11.1	5.6	0.17

Initially based on the notation mentioned above the covariance matrix has been derived from word to word co-occurrences of the financial data which is denoted by 'A'. The probability of the notation is derived using A_{lm} matrix in which the count of occurrence f word 'm' is analyzed based

on the context on occurrence of word 'l'. Let $A_l = \sum_m A_{lm}$ is considered to evaluate the count of occurrence of the word 'l'.

$$P_{lm} = P(m/i) = A_{lm}/A_l \qquad (1)$$

Equation (1) is used to evaluate the probability of occurrence related to word m that appears within the context of word 'l'.

Based on the above scenario the probability of analyzing word to word concurrences is demonstrated by considering the notation depicted in table 1.

Consider two words that include 'l' and 'm' which notates global finance and business news respectively. The co-relational analysis of these words and their co-occurrences could be analyzed by using above probabilistic analysis.

Simple arithmetic is generated between vectors for purpose of meaning among words by performing the difference between vectors we is comparing the words. Infrequent terms lead to unreliable. So, adding weight to frequent terms , we can eliminate infrequent terms

$$\sum_{lm} weight(P_{lm})(dot(A_l, A_m) + a_l + b_m - g(P_{lm}))^2) \qquad (2)$$

The outputs that are generated is given to the canonical correlation analysis approach as a tool to investigate the relationships between the concepts of investment safety and its underlying predicting factors.

Canonical correlation analysis [14] is a statistical model to analyze linear relations among different sets of variables consider to evaluate the percentage of correlation among the current set of variables 'A' and the variables collected from historical data sets 'B'.

Let us assume that the variable set 'A' includes the values from $a_1, a_2, a_3 ..... a_n$ and simultaneously the variable set 'B' includes the values from $b_1, b_2, b_3 .... b_n$ .The canonical correlation analysis among the variables 'A' and 'B' as follows :

Algorithm 1: Canonical correlation among variables 'A' and 'B'.

Input: Canonical Variables from set $A(a_1, a_2, a_3 ..... a_n)$ and $B(b_1, b_2, b_3 .... b_n)$.

Output: Correlation analysis among variable set 'A' and 'B'.

Formation of correlation matrix 'CoM[m][n]' with an order of (P + Q).

i) $CoM_{m,n} = cov(m,n) / t_{mn}$

ii) $Cov(m,n) = \sum_{i=1}^{K} (m_i - m)(n_i - n)/k$

iii) $t_{mn} = \sum_{i=0}^{k} [(m_i - m) * (n_i - n)/2]^2$

Where m and n are the mean values to be calculated from the respective variable sets.

C Processing layer:

After evaluating the correlation analysis among the variables 'A' and 'B' the feasible correlation set of canonical variables are processed within the processing layer which includes Gated Recurrent Unit (GRU) [15] which acts as an filter of unnecessary data where it is notated as generator and

further classification is made based on Convalutional neural networks (CNN) .

D. Gated recurrent unit:

GRU[15] is one of the category of recurrent neural networks i.e., it comes under the context of inter neuronal connection analysis such that which exploits the gradients in traditional RNN based on the long term dependencies.

The proposed GRU includes two gates A_t and B_t. A reset gate R_{st} is used to control the flow of B_t as it is used to analyse previous historical data. As the correlation between the variable set A_t and B_t the influence of B_t on the data of A_t is consider to be negligible if the reset gate are R_{st} is opened.
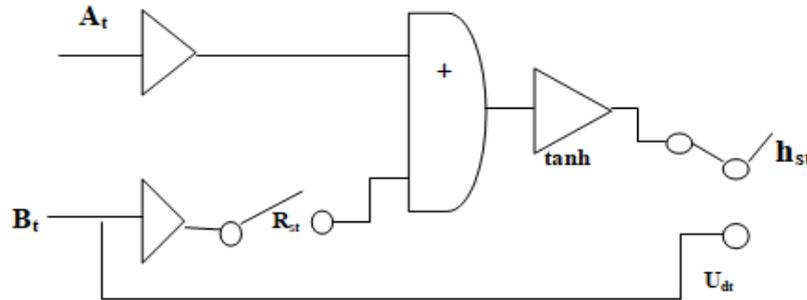

Figure 2 Gated Recurrent Unit

The update gate U_{dt} is used to analyse the current information whether to be ignored or to be consider for the purpose of classification .Simultaneously when the update gate is turned on the uncertain data from the current information set A_t is ignored further to obtain the threshold value.

E. CNN:

Usage of CNN [11] as discriminator within the proposed model will fetch an higher advantage in the context of classification. Generally "conv2D" function could be directly implemented by using Google tensor flow and its application area could be majorly referred for feature extraction and classification of two dimensional images but in our context as we are using stock market data this powerful CNN ability could be used in one-dimensional model for the extraction of patterns within the stock market fluctuations and classify the stock data .Another major factor for selecting 1D CNN as the discriminator ,it is evidently provided that 1D CNNs perform well on spatial data points that are relatively close to each other . On using one dimensional convolution function a CNN model could be generated for prediction as well as classification in the context of classification we will be training the model initially with historic data the working model of CNN as follows initially the vectors of the stock market obtained from the generator module will be given as an input to the stock prediction model where the convolution is computed based on the one dimensional CNN function as one dimensional CNN is considered to be linear function it performs very poor in the context of non-linear problems in such cases we need to implement accelerator function which could also be notated as activation function that includes ReLu sigmoid LreLu and tanh function . In context of solving gradient vanishing problem sigmoid and tanh . The process of CNN is accomplished through the following phases in which initially it is focussed on reducing the computation time of input vector data such that the base of training search data multiple vectors has to be processed in this scenario the input vectors could be represented in the following vector matrix.

$$D= (d^{(1)}, d^{(2)}, d^{(3)}…..,d^{(m)}) \qquad (4)$$

Here D indicated the vector sample in the form of input data that consist of multiple features that are mentioned in the vector matrix such that if in case there are six features matrix may be in the format of L x 6 .the output sequence will be as similar to the input sequence as shown in equation 2.

$$E = (e^{(1)}, e^{(2)}, e^{(3)},…….. e^{(m)}) \qquad (5)$$

In the above equation the e is considered to be as Boolean value that it could be either 0 or 1 as there is only one training sample with one label. In the context of enhancing the convergence speed and to identify the optimal sequence of variable gradient descent technique is used. As there is large variation between the features of the input data it should be normalized by using the following normalization formulae using min-max normalization value within range of 0 and 1.

$$D^{*} = \frac{D-min(D)}{max(D)-min(D)} \qquad (6)$$

As the gradient descent is used there may be chances of gradient vanishing hence for the following activation functions ReLU and Lrelu are used to mitigate the effect of gradient vanishing.

$$ReLu= \begin{cases} 0 \ (D \leq 0) \\ D \ (D \geq 0) \end{cases} \qquad (7)$$

$$Lrelu= \begin{cases} a_i D_i (X \leq 0) \\ D_i \ (X \geq 0) \end{cases} \qquad (8)$$

In the above equation D and D_i are considered to be the outputs generated from the computation of convolution layer of a_i which is assumed as a hyper parameter with minimal initial value (0.01 as initial value) which could be fine tuned during the process of training. Initially in the first ReLu [11] is used to accelerate the process for the Lrelu [11] is used for the next Convolution layer as feature maps with the convolution layer are considered to be input for the polling layer where the reduction of the feature dimensions and there acceleration are simultaneously computed.

At the end of computation process bind with polling layer and convolution layer and significant output layer will be considered as input for the fully connected layer. Where the process of classification is accomplished. Further the softmax function is used in the fully connected layer as shown below.

$$E_i = \frac{f^{c_i}}{\sum_{k=1}^{g} f^{c_k}} \quad i = 2 \qquad (9)$$

Where i represent the number of classification 'C' is the output of fully connected layer and g is the number of C and f is in natural algorithm. Secondly, error function will be computed as follows:

$$H(J^I) = -\sum_{k=1}^{n} \sum_{i=1}^{g} l_{ki}\, log(E_{ki}) \qquad (10)$$

Where J^I is the value obtained from gradient descent step. N is learning rate set to 0.0001.

The classification patterns obtained from the CNN model are further processed using an optimization algorithm to generate accurate predications based on the past data patterns that are trained to the framework.

IV CONCLUSION

This paper proposes an enhanced framework based on generative adversarial network that is integrated with gated recurrent unit as a generator and CNN as a discriminator. Moreover in the initial cases the proposed model uses an NLP tool called Glove for analysing the information from online financial sited where it is processed based on the canonical correlational analysis to identify percentage similar patterns in the data. The studies indicated in the article and the theoretical analysis presented on the model justify that the proposed adversarial network model outperforms any other neural network model in the context of prediction.

REFERENCES

1. Jefferey Pennington, Richard Socher, Christopher D.Manning, GloVe :Global Vectors for Word Representation, Department, Stanford University, Stanford, CA 94305 jpennin@stanford.edu, richard@socher.org, manning@stanford.edu.
2. Jou-Fan Chen, Wei-Lun Chen, Chun-Ping Huang Institute of Information Management National Chiao Tung University HsinChu, Taiwan, R.O.C. , Financial Time-series Data Analysis using Deep Convolutional Neural Networks , 2016, 7th International Conference on cloud computing.
3. Spandan Ghose Chowdhury , Soham Routh , Satyajit Chakrabarti Department of Computer Science and Engineering Institute of Engineering & Management Kolkata, India , "News Analytics and Sentiment Analysis to Predict Stock Price Trends". 2014, ISSN: 0975-9646.
4. Rodrigo Loureiro Malacarne , "Canonical Correlation Analysis ",The mathematica Journal 16 @2014 Wolfram Media.
5. 5R.K. Dase , D. D. Pawar and D.S. Dasput,"Methodologies for Prediction of Stock Market: An Artificial Neural Network E-ISSN:2249-8605, Volume 1, issue 1, 2011 pp 08-15.
6. Rosdyana Mangir Irawan Kusuma , Trang-Thi Ho , Wei-Chun Kao , Yu-Yen Ou and Kai-Lung Hua ,"Using Deep Learning Neural Networks and Candlestick Chart Representation to Predict Stock Market".arXiv:1903.12258v1 26Feb2019.
7. Jithin Eapen Abhishek Verma Doina Bein , "Novel Deep Learning Model with CNN and Bi-Directional LSTM for Improved Stock Market Index Prediction",conference 2019 IEEE.
8. Mustafa A. Sakarwala and Anthony Tanaydin. "Use Advances in Data Science and Computing Power to Invest in Stock Market". SMU Data Scince Review, Volume 2, number 1 Article 17,2019.
9. Mohammad Obaidur Rahman, Md. Sabir Hossain, Ta-Seen Junaid, Md. Shafiul Alam Forhad, Muhammad Kamal Hossen , "Predicting Prices of Stock Market using Gated Recurrent Units (GRUs) Neural Networks".21 march 2019.
10. Koby Todros and Alfred O. Hero III ,"Measure Transformed Canonical Correlation Analysis with Application to Financial Data".IEEE, june,2012.
11. Sheng Chen1,and Hongxiang He, "Stock Prediction Using Convolutional Neural Network", IOP Conf.Series.2018.
12. Xin Lu1 Buyu Li1 Yuxin Yue1 Quanquan Li1 Junjie Yan1 "Grid R-CNN". arXiv:1903.12258v1 29Nov2019.
13. https://in.finance.yahoo.com/.
14. Rodrigo Loureiro Malacarne, Canonical Correlation Analysis, The Mathematica Journal 2014.
15. J. F. Chen, W. L. Chen, C. P. Huang, S. H. Huang and A. P. Chen, "Financial Time-Series DataAnalysis Using Deep Convolutional Neural Networks," 2016 7th International Conference onCloud Computing and Big Data (CCBD), Macau, 2016.

AUTHORS PROFILE

Jasmine Sabeena is Research Scholar in the Department of Computer Science and Engineering at College of Engineering, Sri Venkateswara University, Tirupati, India. She has completed her B.Tech and M.Tech Degree in Computer Science and Engineering. Her research interests are Data Mining .

Dr.P.Venkata Subba Reddy is a Professor of Computer Science and Engineering at College of Engineering, Sri Venkateswara University, Tirupati, India. He is currently the Chairman in the Department of Computer Science and Engineering. His research interests are Fuzzy Inference, Generalized Fuzzy Logic, Fuzzy Dataming. He visited Taiwan to present his papers at iFuzzy 2013, 2014, and 2016.