International Journal of Pure and Applied Mathematics
Volume 118 No. 20 2018, 217-223
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

Predicting Market Prices Using Deep Learning Techniques

Nishanth C P[1], Dr. V K Gopal[2], Vinayakumar R[3], Lakshmi Nambiar[4], Dileep G Menon[5]
[1]PG Scholar, [2]Assistant Professor, [3,4]PhD Scholar, [5]Assistant Professor
[1]Department of Management, [2]Department of Mechanical Engineering, [3]Centre for Computational Engineering and Networking, [4]Department of Computer Science, [5]Department of Management
Amrita Vishwa Vidyapeetham, India
[1]ammasnishanth@gmail.com, [2]vinodkumar@am.amrita.edu, [3]vinayakumarr77@gmail.com, [4]lakshmimavila@gmail.com, [5]dileepgmenon@gmail.com

***Abstract*— This paper discusses the relevance of dynamically changing algorithms to predict the volatile stock market. When an investor decides to buy or sell stock, his decision is highly dependent on the rise or fall of share market price. This study uses a model-independent approach to reveal the hidden dynamics of stock market data using various deep learning techniques such as RNN, LSTM, and GRU. The data from automobile and banking sectors which are listed in NSE is used for the study. Sliding window approach is used for the data analysis and the performance is evaluated using percentage error**

I. Introduction

The benefits of using dynamically changing algorithms to understand the stock market behaviour is discussed in this study with real data analysis. The aim of any investor when he decides to buy or sell shares after analysing the market forecast is profit maximization. Various factors that influence the market behaviour are policy changes, company mergers and acquisitions, global economy, employee layoff, politics, change in management, unmet investor expectation, demand supply behaviour, competitor response and many others.

The stock market is one of the most vital components of a free-market economy, as it provides companies to assess capital. It facilitates stock brokers to trade securities and company stocks. A share may be bought or sold if it is listed on an exchange. Volatile stock market prediction is definitely one of the most challenging areas among time series predictions.

Recently, the methods of predicting stock markets for investing is gaining momentum. Forecasting stock market returns are gaining attention among various investing communities as it provides better guidance with respect to investing. Predictability is one of the major factors which, the profitability of trading in stock and investing is dependent on. Many expert practitioners and researchers have put-forwarded several models using various technical, fundamental and analytical techniques to give a more or less prediction on the stock market pattern. The fundamental analysis involves an in-depth analysis of the changes in the stock prices in terms of external macroeconomic variables; which focuses on determining the intrinsic value of the share. It assumes that the share price of a particular stock is depended on its intrinsic value and the expected return for the investors. But this expected return is subjected to vary as newer and newer information pertaining to a stock is available in the market which in turn alter the price of the share. Investors looking for long-term investments mainly use fundamental analysis [1]. Technical analysis focuses on using price, volume, and historical chart patterns to predict the future stock movements. This model is a short-term prediction of the share price based on its historical evolution.

Beside these commonly used techniques of prediction, some traditional time series forecasting tools are also available. In the time series forecasting, the past data of the prediction variable is analysed and modelled in order to accurately record the patterns of historic changes in the variable. These models are then used for forecasting the future prices. There are mainly two approaches in the time series modelling and forecasting: linear approach and nonlinear approach. Some of the existing linear models are Auto Regression (AR), Auto Regression Moving Average (ARMA), Auto Regression Integrated Moving Average (ARIMA) and its different variants [2]. These techniques are static mathematical models which use predefined equations; as a result, it fails to evaluate the dynamic market changes. Another drawback of these models is that they fail to correlate the relationship between various sectors as the technique is dealing with univariate time series.

The researchers are still searching for techniques to come up with an accurate solution. With more computational capabilities and resources to handle huge databases, complex machine learning, and deep learning techniques are used to address this issue. Few non-linear models used for predicting non-linear time series data are AutoRegressive Conditional Heteroscedasticity (ARCH), Generalized AutoRegressive Conditional Heteroskedasticity (GARCH), Threshold AutoRegressive (TAR) and Deep learning algorithms [2].

The proposed work is trying to find out the relevance of deep learning techniques in predicting stock market price. Three different algorithms, Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) and Gated

Recurrent Unit (GRU) are used in this work for data analysis. The hidden pattern and dynamics of the data can be identified using deep learning algorithms through a self - learning process. The technique also helps to find the dependency between other sectors; however, this study is focused on finding the interdependency between companies in the same sector. The data obtained from National Stock Exchange is used for the study. Sliding window approach is adopted in this paper with data overlap. Companies from automobile and bank sectors are chosen for the study as there is no direct influence between the two sectors. The dataset contains minute wise data of the companies from the above-mentioned sectors. The obtained result can be used to generalize a model for predicting when minute wise data is given as the input.

The paper is structured as follows: Related work in this area is explained in Section [II], Section [III] discusses the background study, Section [IV] discusses the experiments, evaluation results and discussions can be found in Section [V], conclusion and future scope of the study are detailed in Section [VI].

II. Related Work

The prediction of Standard & Poor's 500 index for intraday directional movements given in financial news titles and a set of technical indicators are taken as input parameters in this study. The deep learning methods such as Convolutional Neural Networks (CNN) and Recurrent Neural Network (RNN) are used to predict the next day stock movement. The results of CNN were better compared to RNN for capturing semantics from the texts and RNN outperformed the other on capturing the context information and modelling complex temporal characteristics for stock market forecasting [3]. Google stock price multimedia data from NASDAQ is used for comparing deep learning algorithm with Artificial Neural Network. In this study 2 PCA + Deep Neural Network (DNN) method is compared with state of the art method, 2-Directional 2-Dimensional Principal Component Analysis + Radial Basis Function Neural Network (RBFNN). The proposed method performed better than the existing method RBFNN with an improved accuracy of 4.8% for Hit Rate with a window size of 20 [4]. The results were compared with the Recurrent Neural Network (RNN) and it is found that the accuracy of Hit Rate is improved by 15.6%. The correlation coefficient between the actual and predicted return for DNN is 17.1% more than RBFNN and it is 43.4% better than RNN. Instead of trying to fit to a specific model, the study focuses on the latent dynamics in the data from IT and Pharma sectors using different deep learning methods. The price prediction of NSE listed companies are done using RNN, CNN and LSTM algorithms and their performance is quantified using percentage error. Sliding window approach is used for future prediction in this work. CNN performed better than the other two models in the study [5].

III. Background

The evolution from a rule-based system to deep learning technique is described well in Figure 1. The rule-based system /expert systems are designed by programmers. The required input for the programs is given by the experts in the field based on the application; which includes facts and logic to combine the information into a reasonable solution. In traditional machine learning, the important input parameters are manually designed and the system is trained to map the features automatically. The next step was finding a way to get rid of hand design requirements, which is achieved through representative learning. The main features are automatically learned from the raw data. The deep learning is a high-end version of representation learning. The abstract features in each level are discovered using inputs received from previous levels; thus increasing the level of abstraction. This type of learning is mainly useful for discovering and representing data with high level of abstraction. Thus, deep learning allows learning higher levels of abstractions to unfold the variations and provide easier methods for generalization [6].

Three different deep learning algorithm discussed in this paper are RNN, LSTM, and GRU.

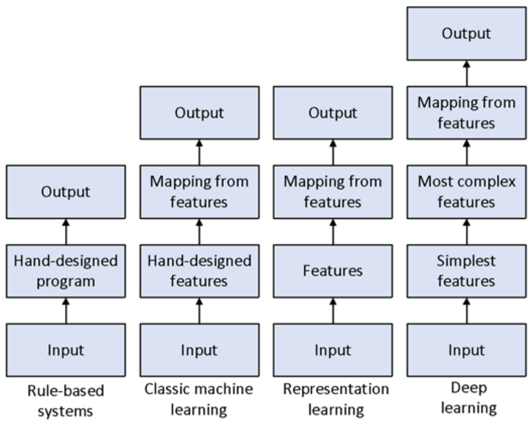


Figure 1: Types of learning methods

A. *Recurrent Neural Network (RNN)*

The Recurrent Neural Networks have feedback loops which can be related to memory. These recurrent connections allow the network to hold information across inputs. It is mainly used to process arbitrary sequence of inputs; particularly useful for learning sequential data like music, stock markets, government agencies, handwriting, spoken words and various other numerical time series data. The way RNN analyses a situation is very similar to real life. The network has two input sources, current, and past data; which is used to determine how interpretations are made when new data is loaded into the network. The feedback loop of the network connects it to the past decisions thus ingesting the moment to moment output of the past as input to evaluate the present data. It also uses the memory for decision making. The hidden state of the network preserves the sequential information. As the feedback loop occurs at every time step in series, each hidden state of the network contains traces of previously hidden state and preceded ones as long as memory exists. The feedback loop allows the information to be passed from one step of the network to another [7]. In many applications long-term dependencies are highly relevant, RNN fails to provide good results. In order to overcome such situation, Long Short-Term Memory technique is introduced.

B. Long Short-Term Memory (LSTM)

This technique introduces a new structure called memory cell. This neural network consists of the sigmoid layer, tanh layer, and point wise multiplication operation. The various components and their functions of this network are: input gate, cell state, forget state, output gate, sigmoid layer and tanh layer. The weight of the self- recurrent connection is given a value ‘1’ to avoid any outside interference and to make sure that the state of the cell remains constant for each time step. The input gate allows the incoming signal to alter the state of the memory cell or block it. The output gate allows the state of the memory cell to have an effect on other neurons or prevent it. The forget gate can modulate the memory cell’s self-recurrent connection allowing the cell to remember or forget the previous state. The cell state runs through the entire network and has the ability to add or remove information with the help of gates. The sigmoid layer generates numbers between zero and one, describing how much of each component should be let through. tanh layer generates a new vector, which will be added to the state [8] .

C. Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) is a version of LSTM. Instead of the three gates as in LSTM, GRU has two gates called update gate and reset gate. It fully writes the contents from its memory cell to the larger net at each time step. The update gate controls the size of previous memory to be kept for data prediction, and the reset gate defines how to combine the new input with the previous memory. Both GRU and LSTM have comparable performance, but GRU is faster to train and requires less data for generalization. If there is enough data, LSTM gives better performance [9].

IV. EXPERIMENTS

A. Proposed Architecture

Deep learning techniques do not rely on any feature engineering mechanisms like traditional machine learning classifiers. Raw data is passed into one or more hidden layers to obtain the optimal features. The architecture used in this work is given in Figure 2, which comprises of four components.

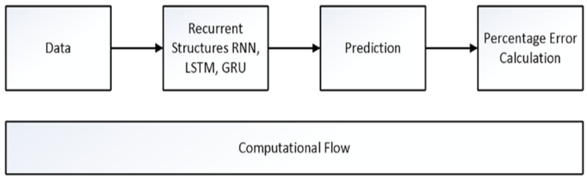


Figure 2: Proposed Model

Data: The study used the minute wise stock price dataset for 1721 NSE listed companies for the trading days from 1st July 2014 to 30th June 2015. The various parameters of the data are day stamp, time stamp, transaction id, stock name, stock price and volume of stock sold in each minute. The two different sectors used in this study are automobile and banking. Tata Motors, Tech Mahindra, and TVS Motors are the companies chosen in the automobile sector and Syndicate Bank, Union Bank and Yes Bank were selected from the banking sector. From the available database, the required sample from each sector is extracted and pre-processed to obtain the stock price. Here we are doing a short-term future prediction, which can be effectively done using sliding window approach. After calculating the error for various window sizes; in this study, it was fixed to be 100 minutes with an overlap of 90 minutes information and prediction was made for 10 minutes in future. In the automobile sector, Tata Motors is used for training which is then tested on the other two companies. In banking scenario, Syndicate Bank is used to predict the performance of the other two banks.

The normalization process is done on the dataset to unify the data to a range of 0 to 1, which was used for training the network. 100 dimension data is passed as an input to the recurrent structures.

Recurrent Structures: The three different neural network structures used in this study are RNN, LSTM, and GRU. Each recurrent structure contains 64 units/memory blocks. The study focuses on finding the existence of long-term dependencies between the given data. The three architectures are capable of identifying the long-term dependencies and the models can be used for future prediction. The interdependencies among various companies can be tested using the model which will help to understand the market dynamics. The train and test data were normalized using the same method.

Regularization: In order to prevent over fitting, the Dropout value was chosen as 0.01, for the recurrent structures. It was also observed that without regularization, the model can easily over fit the training data, even when training on millions of samples.

Classification: Once the features were extracted, the standard dense neural network is used for prediction. The dense neural network has two layers, a Dense (l) unit, followed by the DenseLinear (l) layer with l = 10 units. The dense layers learn a non-linear kernel given the recurrent structures features, and the 'linear' layer output provides the 10-time step predicted values given the output of the final dense layer.

The models were trained for 1000 epoch by varying the layer size to improve the efficiency. If the Mean Square Error (MSE) for the current epoch is less than the value obtained in the previous epoch, the weight matrices for that epoch is stored. After the training process each of the models was tested and the model with the least Root Mean Square Error (RMSE) is taken as the final model for prediction.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Yi - \hat{Y}i)^{2} \qquad (1)$$

In equation (1), Y represents the vector of n predictions, Ŷ represents the vector of n observed values of the variable being predicted. This is an easy method for computing the accuracy of a particular sample. The optimizer used in the study is RMSPROP. Once the predicted output is obtained, de-normalization was applied to the data. Using the available true labels, the percentage error is calculated as given in equation (2)

$$Percentage\ Error = abs\left(\frac{Approximate\ Value - Exact\ Value)}{Exact\ Value}\right) * 100$$

B. Identifying the hyper parameters of recurrent structures

The recurrent neural approaches, RNN, LSTM, and GRU have parameterized functions which will have a direct impact on the optimal parameters of the data thus giving a good prediction. Various experiments and configurations are applied on the data to find out the optimal values of parameters such as learning rate and number of units/memory blocks.

The experiment was initiated by using moderately sized RNN/LSTM/GRU network containing input, hidden recurrent and output layers. Hundred neurons are present in the input layer, the hidden layer contains four units per memory blocks containing one cell each and ten neurons in the output layer. A fully connected network: the connection between neurons in the input to hidden layers and from hidden layer to output layer; is used in the study. Both the training and test dataset values are normalized for the study. Memory blocks in the range from four to sixty-four is experimented and two trial runs are conducted to ensure accuracy. Various other parameters used in the study are: mean square error (mse), as the loss function; RMSPROP, as the optimizer; batch size is 32; data dimension as 100 and a total number of epochs, 1000.

When the algorithms are run with 64 memory blocks, it gave good performance compared to other trials. The performance was also tested by increasing the number of memory blocks from 64 to 128 in each of the algorithms; it was found that the result was similar to the output with 64 blocks. In order to settle on a particular result, the RNN/LSTM/GRU network with 64 memory blocks were run for 1000 epochs.

In order to find the suitable learning rate, two trial runs of the experiments are conducted with different learning rates in the range from .01 to 0.5. Lower learning rate gave better prediction rate though it required more number of epochs before giving final output. The study was conducted with 0.1 learning rate considering various other factors such as training time, cost and prediction rate.

V. Evaluation Results

Three different algorithms used for the study are RNN, LSTM, and GRU. The maximum value of error percentage obtained for each of the models for both automobile and banking sector is given in Table 1 and Table 2. LSTM performed better compared to the other two models. This explains how much the data is dependent on previous value to predict the future instances.

The four test cases used for the study are as follows.

1) Tata motors is used for training the model. It is tested on the test dataset.

2) Tata Motors is used to predict the behaviour of other companies in the same sector.

3) Syndicate Bank is used for training the model and is tested on the test dataset.

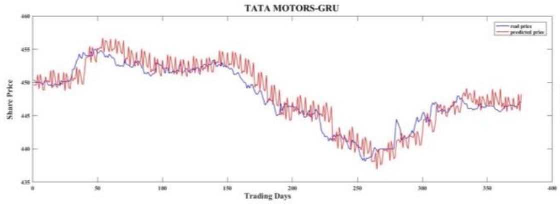
4) Syndicate Bank is used to predict the behaviour on other banks.

Company	**RNN**	**LSTM**	**GRU**
Tata Motors	1.8449	1.4747	1.7346
Tech Mahindra	7.1415	6.24721	6.6558
TVS Motors	2.0057	1.7926	2.0580

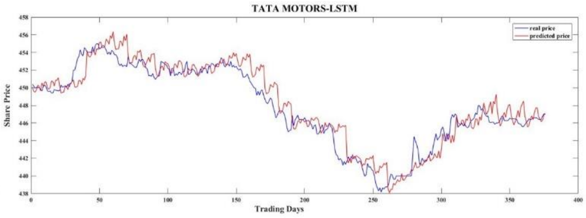
Table 1: Error Percentage for Automobile Sector

Bank	RNN	LSTM	GRU
Syndicate	1.7233	1.5568	1.6481
Union	1.5740	1.3485	1.4062
Yes	1.6709	1.7281	1.6958

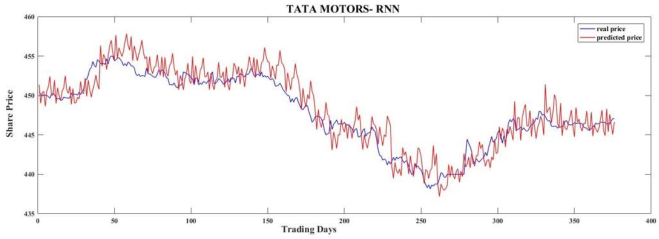
Table 2: Error Percentage for Banking Sector



(a)



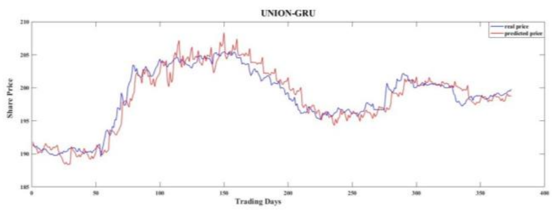
(b)



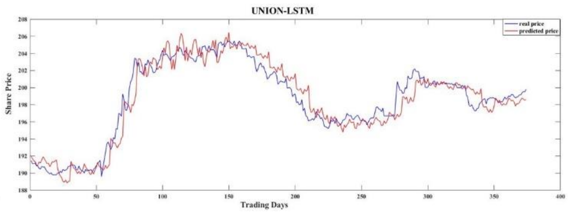
(c)

Fig 1: Plot of Real Vs Predicted Value of Tata Motors using GRU, LSTM and RNN

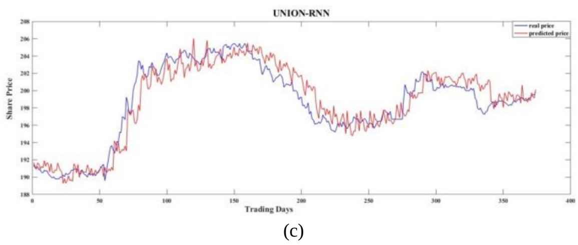
Figure 1 (a), (b) and (c) gives the details of the prediction result for the three algorithms GRU, LSTM, and RNN respectively for Tata Motors. The measurements are based on shared price and trading days. The variation in the actual and predicted is higher for RNN compared to the other two. LSTM gave better prediction result for all the three. Initially, the prediction result was very much in alignment with the actual output. The time period between 100 and 150 trading days also gave a good result. There was an unexpected market change between 150 and 180 time period; which was reflected by a difference in actual and predicted values. This condition was learned by the algorithm and the prediction result between 180 and 220 time period gave an improved result before a major shift occurred around time period 230. We can also see from the plot, figure (b) that the algorithm is getting improved as it proceeds. Around 370 trading days, the actual and predicted values are totally in alignment.



(a)



(b)



(c)

Fig 2: Plot of Real Vs Predicted Values of Union Bank using GRU, LSTM and RNN

Figure 2 (a), (b) and (c) gives the plot results of predicted and the real price of Union bank. As seen in the automobile sector, LSTM algorithm performed better than the other two algorithms. The variation was higher for RNN. The data was highly fluctuating; especially between the time period 50 - 100. Although the algorithm failed to give the exact value, it gave a tentative result compared to the other two. Between the time period 300 - 350, the prediction result was much better. GRU also gave good predictive result during the same time period.

As we know it is very difficult to give an accurate prediction for volatile data, Long Short-Term technique gave a better result compared to other algorithms in two sectors. It can also be observed that even though LSTM performed better than GRU, the variation is minimal.

VI. Conclusion and Future scope of the study

Deep learning technique is slowly emerging as a promising tool for predicting non-linear time series data like the stock market. The hidden dynamics and abstraction of the data can be found out to a certain extent by using this technique; thus giving a more accurate prediction in comparison to many existing models. In the study, data from two sectors were used for analysis. In the automobile sector, data of Tata Motors is used for training and predicting the stock price of Tata Motors, Tech Mahindra, and TVS Motors. In the banking sector, Syndicate bank data is used for training and it is tested on Union Bank, Yes Bank and Syndicate Bank using various algorithms such as RNN, LSTM, and GRU. The technique was able to identify interrelation between various companies within the same sector; the best result was obtained using LSTM model. The trends and periods of existence will vary according to the sector and companies chosen. The ultimate aim of the study is to analyse the data and predict the different trends and cycles that will give more profit to the investors. The study can be further extended by using techniques such as Convolution Neural Network (CNN) to capture the irregular changes occurring in the stock market. The relationship between various sectors can also be evaluated thus allowing us to find out if there are any hidden parameters which will correlate the performance of various sectors that appear totally independent from the first glance.

Acknowledgement

I take this opportunity to express my sincere thanks to Dr. Beena S Nair, Vice Chairperson, English Department for proofreading this work and for her valuable comments. My sincere thanks to Mr. Vinu Krishnan, M.Tech student for his valuable time in helping me get the diagrams together. My sincere thanks to Dr. K.P Soman, H.O.D, Computational Engineering and Networking (CEN) for giving me all the support and expertise. I also wish to extend my token of appreciation to Mr. R. Pradeep, PhD Scholar, EEE Department of ASE for his support and encouragement. My gratitude to all my teachers, friends, family and my guru, Her Holiness Mata Amritanandamayi Devi, Chancellor, Amrita University, for her grace throughout the work.

References

[1] Virginica, Rasu & Cristina, Rasu (2003). Forecasting methods and stock market analysis.

[2] Manna, Majumder & MD Anwar Hussain. Forecasting of Indian stock market index using artificial neural network.

[3] Manuel R. Vargas, Beatriz S. L. P. de Lima, Alexandre G. Evsukoff (2017). Deep learning for stock market prediction from financial news articles: IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (2377-9322)

[4] Ritika Singh, Shashi Srivastava (2017). Stock Prediction Using Deep Learning: Journal of Multimedia Tools and Application, Volume 76 Issue 18, September 2017, Pages 18569-18584

[5] Sreelekshmy Selvin, Vinayakumar R, Gopalakrishnan E.A, Vijay Krishna Menon, K P Soman."Stock Price Prediction Using RNN, LSTM and CNN Models – Sliding Window Model.", International Conference on Advances in Computing, Communications and Informatics (ICACCI)

[6] Yoshua Bengio, "Deep Learning: Theoretical Motivation".Video lecture given at Deep Learning Summer School, Montreal Canada in August 2015

[7] Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, *78*(10), 1550-1560.

[8] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

[9] Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, *5*(2), 157-166.

[10] Ian Goodfellow and Yoshua Bengio and Aaron Courville. (2016). Deep Learning. Book, Goodfellow-et-al-2016,MIT Press

[11] Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179-211.

[12] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

[13] Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM.

[14] Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. *Journal of machine learning research*, *3*(Aug), 115-143.

[15] Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (pp. 315-323).

[16] Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In Proc. ICML (Vol. 30, No. 1).