Time Series Forecasting using an ensemble of Artificial Neural Networks from a cluster of Neural Nets

Tushar Agarwal
ta2482@columbia.edu
Dept. of Computer Science, Columbia University

***Abstract*— Artificial Neural Networks(ANNs) have found applications in many areas of finance and trading. The use of ANNs to solve Time Series Forecasting problems is becoming quite popular. However, one problem faced by traders is that with time, an existing Neural Network Model may become obsolete. Preparing a new model comes at a cost of both computational time as well as model preparation time. In this paper, we propose a Novel scheme that uses a cluster of Neural Network Models to solve this problem.**

***Keywords*— Artificial Neural Networks (ANNs), Time series forecasting, Ensemble, Distributed Clusters**

I. INTRODUCTION

Enhancing the accuracy and robustness of time series forecasting models is an active area of research in the field of computer science. Analyzing and forecasting of time series problems is of fundamental importance in several practical domains. In time series forecasting, the available historical observations are carefully studied in order to build up a proper model. This is then used to forecast unseen future values [1]. Over the past years, various non-linear and linear forecasting models have been developed in academic literature. During the last two decades, Artificial Neural Networks have been widely used as attractive and effective tool for time series forecasting and modeling [7].

ANNs have been originally motivated by the intelligent neural structure of human brain. With time ANNs have found extensive applications in solving a broad range of nonlinear problems and have been drawing increasing attention of the research community. The most distinguishing feature of the ANNs is that they are nonlinear, nonparametric, data-driven and have self- adaptive nature. ANNs do not require any a priori knowledge of the associated statistical data distribution process. They can adaptively construct the appropriate model by using only the raw data, learn from training experiences, and also intelligently generalize the acquired knowledge to predict the outcomes of unseen future events. Because of these outstanding properties, they have become a favorite choice for time series researchers as well as traders in the world of finance. Currently, many ANN based forecasting techniques exist in literature. Some excellent details and review on recent trends and development of ANN forecasting methodology can be found in the works of Kamruzzaman et al. [8], Zhang et al. [9], Adya and Collopy [10] and Felipe Giacomel, Renata Galante and Adriano Pereira [11].

While ANNs have been proven to be quite effective for forecasting any time series problem, a common problem faced by traders is that a particular ANN model may become obsolete with time due to changing behavior of the time series signal. Thus the trader will have to remodel the ANN for future predictions. This paper is focused on providing a solution to this problem. The paper is divided into multiple sections. In section [II] we describe the problem in detail. In section [III], we discuss current existing solutions and their limitations. Section [IV] presents the motivation behind proposed solution whereas section [V] presents the complete solution and the key design considerations. Section [VI] gives an example of stock price prediction using the proposed solution. Section [VII] details the advantages of the solution while section [VIII] highlights the various challenges posed by the proposed solution. Finally, section [IX] concludes the paper.

II. PROBLEM DESCRIPTION

An intelligent trader will focus and spend a lot of time on deciding the governing input items for his or her neural network. Also he or she will require to adjust the parameters of the ANN model to attain best predictions. For this he or she will spend several weeks and sometimes several months deploying the network. However, with time, an existing ANN model may become obsolete due to the changing conditions of a Time Series Forecasting problem and a different model may be required. Existing practices suggest that the trader can solve this issue by retuning the parameters of the existing Neural Network model or sometimes remodeling the whole Neural Network. But all this comes at a cost of both computational time as well as model preparation time. Furthermore, a single neural network may only cover a relatively small aspect of the market and hence many neural networks may also be required to be modelled to achieve a meaningful prediction.

Often it is also difficult to decide the time series interval for which to to model a Neural Network. Notice that the time step of a series may be of any length, for example: seconds, hours, days, years etc. This will bring out very different looks of the time series. For example, Fig. 1 depicts variation of stock price for IBM and Microsoft. The first image in Fig. 1 shows the variation of stock price over a year with time step of 3 months. The later image shows the variation in stock price for a portion of same day with

Fig. 1. The above Figure shows variation of stock price of IBM and Microsoft. The former image depicts variation in stock price over the year 2016 with time intervals of 3 months whereas the later depicts the variation only for a portion of the same day with time interval of 2 hours. [Image Source: Yahoo Finance. Image was taken on 13th December 2016]

step size of 2 hours. Thus a trader may require to re-tune his model to a different step size based on the behavior of the time series signal.

In this paper, we suggest a scheme which does not require the Trader to model a new Neural Network when conditions change for a Time Series Forecasting problem. The scheme also cuts on the computational time required for learning an ANN.

III. Limitations of the Current Solutions

The estimation of future values in a time series is commonly done using past values of the same time series. Given a time series problem, forecasting refers to the process of calculating and predicting one of several values ahead, using just the information given by the past values of the time series. One of the prevalent time series solution using Artificial Neural Networks is Long-Short Time Memory architecture which consists of a set of recurrently connected subnets. The objective of the LSTM architecture is to overcome the problem known as the "vanishing error problem" [6]. The vanishing error problem refers to how the influence of past inputs decays quickly over time. LSTM networks aims to solve this problem using memory cells.

While these kind of neural networks are quite efficient for any time series forecasting problem, they start failing when the time series signal changes its behavior and becomes more dependent on some other segment or factor of the market as discussed in section [II]. This calls for an efficient solution which will not require a trader to remodel his ANN every time the behavior of the time series signal starts changing.

Many approaches also use a static ensemble of neural networks to solve time series forecasting problems [13]. However, again these are not able to solve the problem mentioned in the paper. The static ensemble does not update once the neural networks in it start becoming obsolete.

IV. Motivating the Solution

Imagine yourself to be an instructor taking a class of machine learning comprising of 50 students. Students in this class have enrolled from wide diversity of disciplines. Some are from Computer Science background, some from Mathematics and others from Finance background. You want to develop an Artificial Intelligence model for say

weather prediction. You want to choose 5 students from your class to do this project. Who should you choose? Trivially, you would pick out the 5 top performers in the class and ask them to do the job. But let's say after sometime the requirement of the project changes and the team needs some mathematical problems to be solved. Currently the ensemble of top 5 students you choose did not have any Machine Learning student with Mathematics background. So you pick the best Mathematics background student from your class and add it to the team. You may remove one person from the team to keep the team size constant. This approach is exactly what we propose to do with a cluster of Artificial Neural Networks to solve the problem defined in the paper.

In our solution, we form a distributed cluster of machines. Each machine can train a neural network model. Several neural network models each representing a solution to the given Time series Forecasting are trained on this cluster. An ensemble is later built out of this cluster depending on the models which give best predictions. The ensemble updates the choice of models it uses for prediction depending on the behavior of the time series signal. We formally state the algorithm in the next section.

V. The Solution

The solution to the problem defined in the paper consists of two parts The Before Step and The After step. The former is the step which trains a cluster of Neural networks and chooses an ensemble from it. The later elicits how the ensemble runs and updates while predicting in real time.

A. Algorithm

"The Before Step:"

- The Trader must decide on $N_1, N_2....N_m$ Neural Network models for the Time Series problem he is trying to solve. Each model may represent only a certain aspect of the market or the whole but with different parameters.
- Next create a distributed cluster of machines to train all the Neural Network models. Let's call it C (Fig. 2).

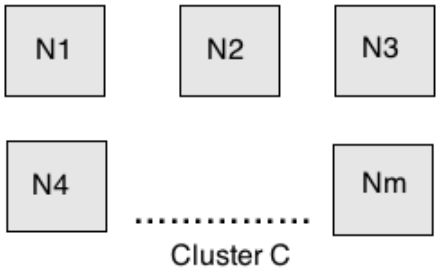

Fig. 2. Cluster C

- Pick the relevant time series data and break it into portions $T_1, T_2, T_3, ...T_n$ as shown in Fig. 3.

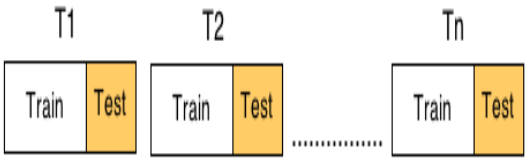

Fig. 3. Breaking Time Series Data

- For all T_i do: Break it into Training and Validations sets. Train all models in C and store the test results corresponding to each T_i.
- Give rank to each model based on its ability to predict over time.
- Create an ensemble of top k ANNs where m $>$ k, based on the ranks allotted in previous step.

"The After Step:"

- Use the above ensemble to get predictions on new Time series data. A voting may be performed to reach a consensus.
- Once sufficient new labeled data is available, pass it again through C. Re-compute ranks of each model. Update the Ensemble if required.

B. Design Considerations

For the success of the proposed scheme, it is important that the trader models the neural networks carefully. Following key constraints must be considered while designing the Neural Network models:

System Architecture:

1) Total number of neural networks to be modeled to build the cluster
2) Number of neural networks to be chosen for the ensemble. Current practices suggest the number should be between 5-10 [12].
3) Different segments of the market which a neural network should reflect.
4) A single Neural Networks parameters and hyper-parameters [4]. These include:
   - The interconnection pattern or the Weight vector between the different layers of neurons
   - The learning process or the training algorithm for updating the weights of the interconnections.
   - The activation function that converts a neurons weighted input to its output activation.
   - Number of input nodes
   - Hidden layers and hidden nodes
   - Output nodes

The above has been discussed in more details in the mid-term paper. A typical Neural Network structure is presented in Fig. 4.

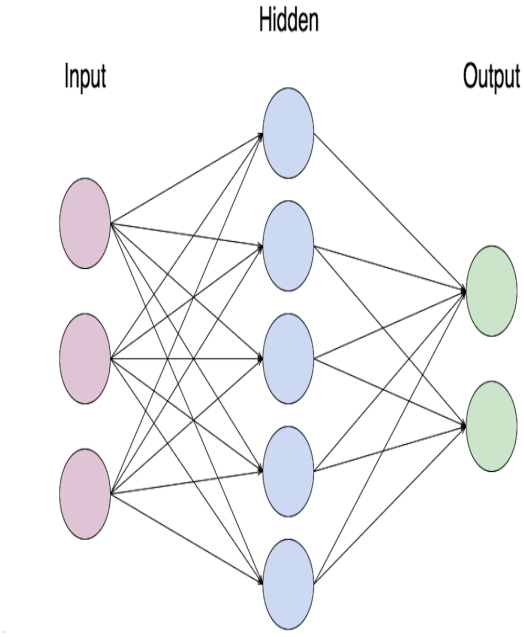

Fig. 4. A typical Neural Network Diagram

Cluster Deployment:

- Another Key design constraint is the formation of the distributed cluster for training several artificial Neural Network models. Forming a huge cluster can be quite expensive and cumbersome. However, such a problem was solved by google back in 2003 by introducing the Google File System(GFS) [3]. GFS now has an open source version introduced by Apache and famously known as Hadoop [14]. Hadoop is an open-source distributed computing framework which can be used to deploy a cluster of machines using cheap commodity hardware. Our ANN cluster can be deployed using a similar scheme. This will enable us to form a rigid cluster of ANNs as well as cut on costs by using cheap commodity hardware.

Ranking Algorithm:

- It is important to decide how we are going to rank the ANN models. One approach is to rank the model which has the highest accuracy throughout time as 1. Rest of the models follow in a similar manner. Another approach can be to give more weight to the model which predicts best using the latest time series data. Thus, this is a hyper-parameter on which the trader must decide.

VI. Time Series Forecasting of Stock Price: An example

Stock price prediction is a famous Time series problem. To achieve good results in stock price prediction using the Novel scheme presented in the paper; the trader must choose a good neural network model or models, for example - a recurrent neural net [2]. Recurrent neural nets are quite suited to exhibit dynamic temporal behavior. They are a perfect choice for problems such as Time Series Forecasting.

The trader must also get as much data as they can. This may include stock price, volatility, correlations, earnings, volume etc. The order book - a catalogue of what sales and buys have been offered at a particular moment is particularly valuable and may be used to model some neural networks in the cluster. Ultimately the trader must apply and run the algorithm as shown in Fig. 5.

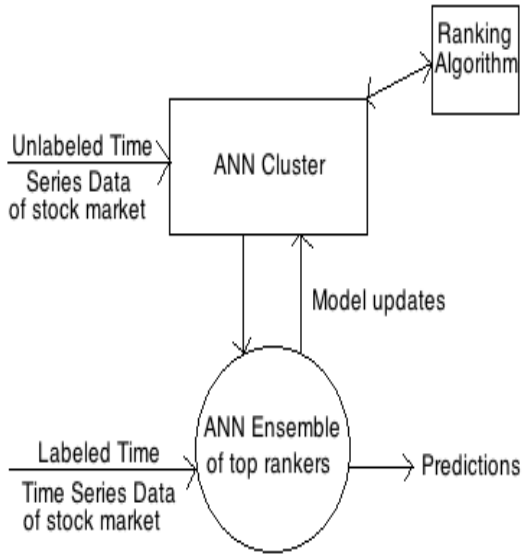

Fig. 5. The above diagram shows the complete architecture as proposed by the algorithm given in the paper. Labeled data is input to ANN cluster which trains m Neural Network models. Then an ANN cluster of top ranker's is formed. New Unlabeled data of stock market is input to the ensemble which makes predictions. Models in the ANN ensemble keep updating as new labeled data is available.

VII. Advantages

The proposed scheme offers various advantages. Some of them are:

- No Need to conceptualize a new model every time behavior of Time series data starts changing.
- From a class of Neural Networks, the top k Intelligent Neural Networks participate in prediction thus improving the chance of predicting correctly while reducing computations.
- A cluster of machines, where each ANN may have a dedicated system or multiple systems helps in achieving faster computational speed.
- Clusters can be built using cheap commodity hardware thus reducing infrastructure costs.

VIII. Challenges

A. Implementation Challenges

The proposed solution does pose challenges for the Trader. Sometimes, it could be difficult for him to decide

an exhaustive set of neural networks to build the cluster from. For example, he may not be able to decide whether he should have a neural network which takes in data at a time interval of minutes, hours, days, months or anything in between. He can thus form infinite such combinations based on his time series problem. Also there may be many latent variables to a time series problem which may not be visible to the trader at first. Thus, the trader must decide on what neural networks to build the cluster from after careful consideration.

Apart from this, there certainly are challenges to implementing a single artificial neural network [4]. These include design decisions such as: - deciding the inputs, number of hidden layers and the number of neurons in each layer, the selection of activation functions of the hidden and output nodes, the training algorithm, data transformation or normalization methods, training and test sets, performance measures etc.

B. Adoption and Regulatory challenges

Neural Networks or any predictive paradigm by nature is often conflict-based, dispute-based or highly sensitive, and the perception that predictive coding introduces uncertainty hinders its adoption [5]. A trader must consider following key points carefully before using ANNs to solve a problem:

1) Determination of the criteria for measuring the precision of a sampling process of a time series problem and how workflow steps improve the ANN.
2) Learn how it all works so that the authenticity of the model can be proved in court if in question.
3) The trader must maintain proper transparency and communication with the customer regarding the choice and deployment of ANNs.

IX. CONCLUSION

The paper presented a novel scheme to solve any Times Series Forecasting problem using Artificial Neural Networks. It solves the issue of forming new ANN models once an existing model becomes obsolete. The scheme claims to be able to reduce computational time by using a distributed cluster of machines. It also promises to give good results because of the ensemble created on the basis of the rank of the classifiers.

REFERENCES

[1] G.E.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, 3rd ed. Holden-Day, California, 1970.

[2] Hochreiter, s., Schmidhuber, J. Long Sort-term Memory. Neural Computation, (1997), 9 (8), 1735-1780

[3] Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung Google, "The Google File System", SOSP03, October 1922, 2003, Bolton Landing, New York, USA.

[4] Tushar Agarwal, Artificial Neural Networks in Finance and Financial Forecasting

[5] Jie Zhang, Hype Cycle for Legal and Regulatory Information Governance, 2016 Gartner Report

[6] INAOE 2014. P. GOMEZ-GIL et al.

[7] Ratnadip Adhikari, R. K. Agrawal, A Homogeneous Ensemble of Artificial Neural Networks for Time Series Forecasting , International Journal of Computer Applications (0975 8887)

[8] J. Kamruzzaman, R. Begg, R. Sarker, Artificial Neural Networks in Finance and Manufacturing, Idea Group Publishing, 2006.

[9] G. Zhang, B.E. Patuwo, M.Y. Hu, Forecasting with artificial neural networks: The state of the art, International Journal of Forecasting 14, pp.3562, 1998.

[10] M. Adya, F. Collopy, How effective are neural networks at forecasting and prediction? A review and evaluation, Journal of Forecasting 17, pp. 481495, 1998.

[11] Felipe Giacomel, Renata Galante,Adriano Pereira, An Algorithmic Trading Agent based on a Neural Network Ensemble: a Case of Study in North American and Brazilian Stock Markets, 2015 IEEE/WIC/ACM International Conference

[12] http://www.investopedia.com/articles/trading/06/neuralnetworks.asp?lgl=bt1tn-above-textnote

[13] V. Landassuri-Moreno, John A. Bullinaria, Neural Network Ensembles for Time Series Forecasting, GECCO09, July 812, 2009, Montral Qubec, Canada.

[14] https://hadoop.apache.org/docs/r2.7.2/api/org/apache/hadoop/fs/FileSystem.html