Artificial Neural Networks in Finance and Financial Forecasting

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

Abstract

An Artificial Neural Network is a computing system modeled after the human brain and the nervous system to process information. Finance and trading are one of the most frequent areas of artificial neural network applications. Time Series Forecasting problems such as stock market predictions, foreign exchange rate prediction, etc. is one area in the domain of finance where artificial neural networks are trending. In this paper we present artificial neural networks approach to predict Time series forecasting problems. We also talk about the scope of artificial neural networks in the financial domain.

***Keywords*—** Artificial Neural Networks (ANNs), Time series forecasting, stock market, algorithmic trading

I. INTRODUCTION

The modern era is witnessing a boom of artificial intelligence and its technologies. Artificial Neural Networks is one such paradigm that has seen an explosion of interest over the last few years. ANNs have found many applications in various domains including finance, medicine, physics, engineering and geology. Many attempts have been made to formally define neural networks.

”Artificial neural systems, or neural networks, are physical cellular systems which can acquire, store, and utilize experiential knowledge.” - Zurada (1992)

”A neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes” - DARPA Neural Network Study (1988)

Artificial Neural Networks came into existence due to various research being conducted in the domain of Artificial Intelligence. When traditional predictive systems based on symbolic searches started to fail, specific attempts to build systems that mimic the architecture of the human brain were made. The fault-tolerance, interconnected network of neurons and learning capacity of biological neural systems helped conceptualize the idea of artificial neural networks.

In this paper, we explain the paradigm behind artificial neural networks and also discuss its applications in financial domain. The paper is divided into multiple sections. Section [II] starts with an overview of the artificial neural networks, followed by the description of the ANN model, its key design considerations, the training algorithm and the activation function. Section [III] describes how neural networks are implemented to solve time series forecasting problems such as stock price prediction, algorithmic trading, etc. Section [IV] gives a brief overview of some other applications of neural Networks in Financial industry. Section [V] describes various advantages and disadvantages of using ANNs. Section [VI] talks about various challenges faced to implement, adopt and regulate ANNs. Section [VII] presents the Key Players who use ANNs to solve Financial problems. It also presents how FINRA uses ANNs for market regulation. Section [VIII] gives a peek into the outlook and future trends surrounding ANNs and its applications in Finance. Finally, section [IX] concludes the paper.

II. ARTIFICIAL NEURAL NETWORKS

An Artificial Neural Network (ANN) is a supervised Machine learning algorithm that has been inspired by the way biological nervous systems such as the human brain functions (See Fig. 1). It is composed of a large number of highly interconnected processing elements called neurons, working in unison to solve specific problems. ANNs, like biological nervous systems, learn by example. An artificial neural network is configured for a very specific application like a data classification or pattern recognition platform through a learning process. Learning in biological systems involves incorporating adjustments to the synaptic connections which exist between the neurons. This is true for the artificial neural networks as well.

A. *Artificial Neural Network Model*

Neural networks are typically organized in layers, the first one being the input layer, the last one is output layer, while all the intermediate ones are called hidden layers. Each layer is made up of a number of interconnected ’nodes’ consisting of an ’activation function’. Data is presented to the network via the ’input layer’, which communicates to one or more ’hidden layers’ where the actual processing is done through a system of weighted ’connections’. Finally, the hidden layers then link to an ’output layer’ where the answer is output (See Fig. 2). The architecture of the network defines how the nodes in a network are interconnected. Data flows through this network in one direction only, i.e. from the input layer to the output layer. Each layer consists of computational

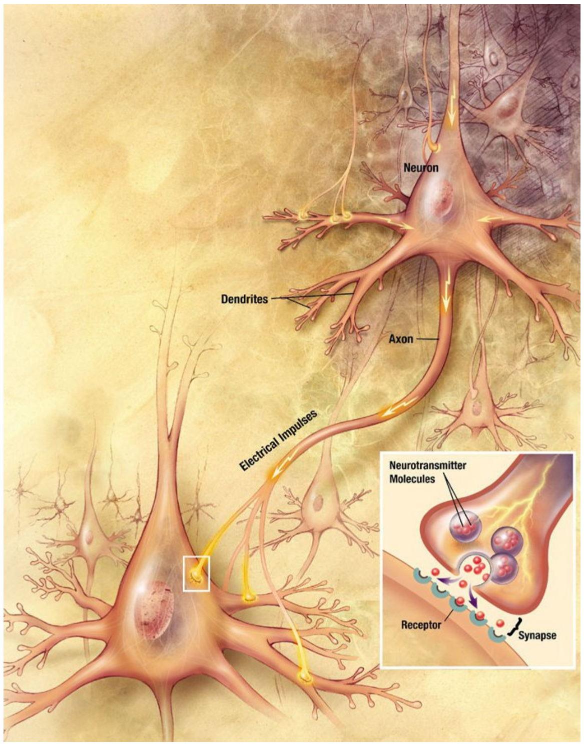

Fig. 1. The above image depicts basic structure of a neuron, with its axons and dendrites indicated. Information flows from one neuron to another at the synapse in form of Electrical Impulses. [Image Credit: National Institute on Aging / National Institutes of Health]

units(neurons) or hidden nodes which receive signals from other computational units and maybe a bias adjustment. The sum value - Σ for any computational unit is the linear combination of each signal received by the unit multiplied by the weight of the corresponding connection. The output y_i for this unit, where i represents the layer number, is the result of applying a transfer function g() to the sum value Σ.

$$\Sigma = W_0X_0 + W_1X_1 + W_2X_2 + ... + W_nX_n$$
$$y_i = g(\Sigma)$$

B. Key Design Considerations

An ANN is typically defined by three types of parameters:

1) The interconnection pattern or the Weight vector between the different layers of neurons (W in Fig. 2).
2) The learning process or the training algorithm for updating the weights of the interconnections.
3) The activation function (g() in Fig. 2) that converts a neuron's weighted input to its output activation.

While building ANN, one must also determine the following variables:

- Number of input nodes (X_1, X_2 ... X_4 in Fig. 2)
- Hidden layers and hidden nodes (layer 2 and layer 3 in Fig. 2)
- Output nodes (y in Fig. 2)

The selection of the above parameters is basically problem-dependent.

C. The Training Algorithm

The artificial neural network training is an unconstrained nonlinear minimization problem in which weights of different arcs in the network are iteratively modified to minimize the overall error on some Training Data. There exist many different optimization methods that provide various choices for neural network training. The most popularly used training method is the back propagation algorithm [16]. A back propagation neural Network uses a feed-forward topology, supervised learning, and the back propagation algorithm. It mainly employs some form of gradient descent algorithm [18], using back propagation to compute the weight vector. Another variation known as the Recurrent back propagation [17] is a network with feedback or recurrent connections. Addition of recurrent connections to a back propagation network intensifies its ability to learn temporal sequences without fundamentally modifying the training process. It therefore performs even better than the regular back propagation network.

D. The Activation Function

This function is responsible for determining the output of the input values in a node in an artificial neural network. This is the function which is responsible for introducing non-linearity to an ANN. Some examples of activation functions are:

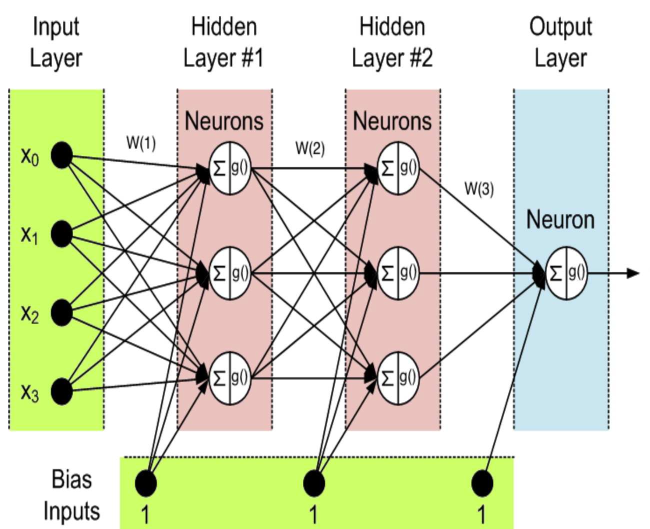

Fig. 2. A typical representation of a Neural Network. Here X_0, X_1, X_2, X_3 are the inputs; W(1), W(2), W(3) are weight vectors; g() represents the transformation function; 1 is present as the bias in above example; Σ represents summation of multiplication of weight vector W with the corresponding input signal vectors.

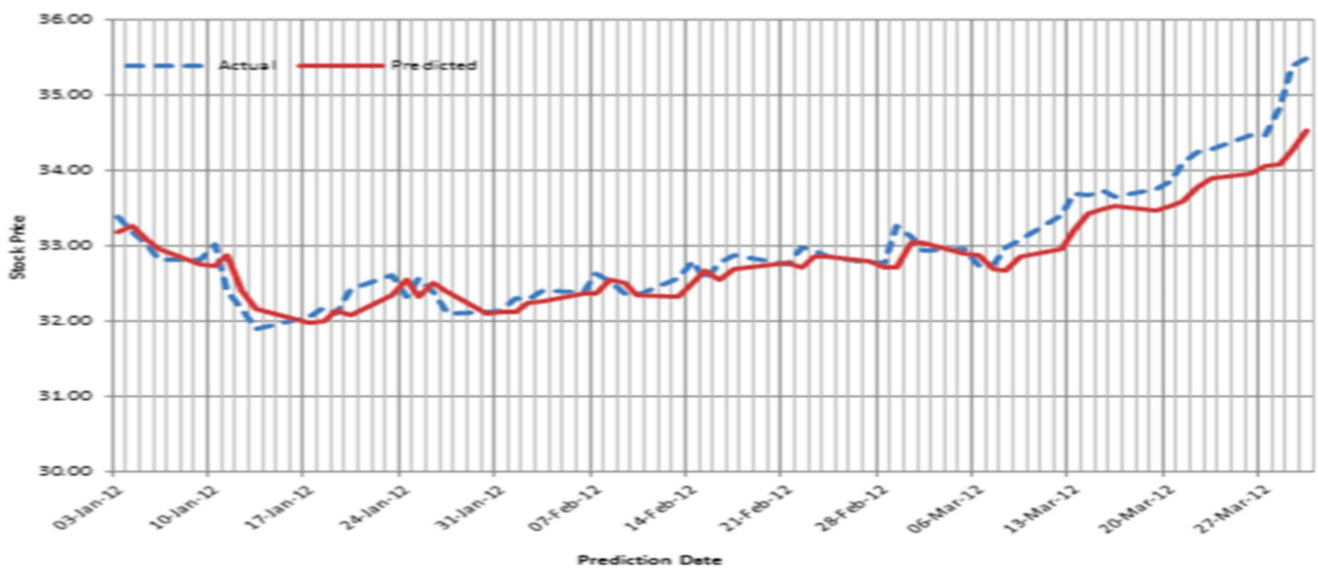

Fig. 3. Coca Cola Stock price: Predicted vs Actual price using ANN, Image source [2]

1) The Logistic/Sigmoid Function: $f(x) = \frac{1}{1+\exp(-x)}$
2) Sine or Cosine function: $f(x) = sin(x)$ or $f(x) = cos(x)$
3) Hyperbolic tangent function: $f(x) = tanh(x)$

III. APPLICATION OF NEURAL NETWORKS IN TIME SERIES FORECASTING

Research and studies have shown the power of the artificial neural networks for classification and prediction problems. It has been well researched and demonstrated that a neural network can approximate any continuous function. Time series analysis requires the user to analyze continuous stream of data in order to extract meaningful statistics and other characteristics from it. Time series forecasting uses the knowledge derived from the time series analysis to form a model which can predict future values based on previously observed values.

Neural networks have been successfully used for time series forecasting of financial data. Some of the time series forecasting problems in finance industry are:

1) Stock market predictions,
2) Foreign exchange rate,
3) Algorithmic Trading,
4) Gold price,
5) Financial prognoses (returns on investments)
6) Commodity prices, etc.,

The classical methods used for time series prediction like Box-Jenkins, Auto-regression and Auto-regression Integrated Moving Average (ARIMA)[1] assume that there is a linear relationship between inputs and outputs. Whereas, artificial neural networks have the advantage of being able to approximate any nonlinear functions without any prior information about the properties of the data.

For example, in stock markets, investment is usually guided by some form of prediction. To start with, there is need to model the trend of the prices of the stock, which is nonlinear. An important characteristic of stock market data is the dependence on time, hence current observations depend on past observations. Since ANN can model a continuous function, it is the best choice for prediction. Fig. 3 shows us actual vs predicted stock price of Coca Cola over time using ANNs.

Another example is of algorithmic trading which uses algorithms to drive trading decisions, usually in electronic financial markets. Applied in buy-side and sell-side institutions, algorithmic trading forms the basis for high-frequency trading, foreign exchange trading, and also its associated risks and execution analytics. In Algorithmic trading, the traders do not participate in trading (either buying or selling of shares) if there is no substantial change in the price of the security being traded.

Many algorithmic trading applications need to develop trading strategies using non-linear time series machine learning methods. This is where deep learning and artificial neural networks become useful.

Now that we know how artificial neural networks are useful to solve a times Series forecasting problem, lets have a look at how do we apply ANNs to a time series problem. The building of a neural network for time series forecasting goes through the definition of following steps:

1) Identification of the forecasting target
2) Building the data set upon which to activate neural network learning
3) Activating the network for learning

A. Identification of the Forecasting Target

The first step before starting to build a neural network for financial applications involves identifying the forecasting target. The applications of neural networks are quite diverse. It may be for stock market prediction, foreign exchange prediction, predicting the gold price, etc. It is necessary to establish the borders of the measurement of the phenomenon to be evaluated. For example, if we want to predict how the price of gold will vary in the coming few days, weeks and years, it is necessary to consider all the factors which impact the price of gold in our model.

B. Building the data set on which to activate ANN learning

The next step involves:

1) Collection of Data: Information gathering must meet some fundamental principles. For example, in stock market prediction, data must be recovered from markets regularly in order to guarantee the historical series continuity.
2) Analysis and Transformation of Data: This process tends to identify the optimal data sets depending on the analyzed problem, according to an iterative scheme that must guarantee the maximum learning ability with the minimal informative effort. For example, an optimal set may be said to have data tuples at least 5 times as large as the number of input variables.
3) Choosing the Input and Output Variables: The model needs to consider only those data features, which will help correctly classify the output variables in question. For example, for stock price prediction of Goldman Sachs, the year in which company started may not be an important input variable, whereas historical variation in the price of stock of Goldman Sachs will play a very important role.

C. Activating the ANN for learning

This step involves the analyst to choose the architecture and parameters necessary for the definition of the connection weights between neurons as discussed in section II-B of the paper.

Once that is done, a training and a test sample are typically required for building an ANN forecaster (Fig. 4). The training sample is used to build the ANN predictor and the test sample is used for evaluating the forecasting ability of the model. Various literature state that data may be divided into training and test sample based on the rule of 90% vs. 10%, 80% vs. 20% or 70% vs. 30%, etc. Once the ANN model is ready, it can be used predict the underlying time series forecasting problem.

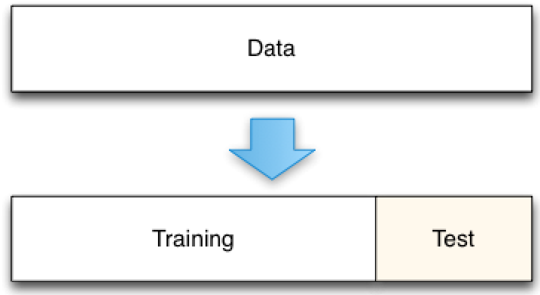

Fig. 4. Data divided into Training and Test to validate the model

IV. SOME OTHER APPLICATIONS OF NEURAL NETWORKS

A. Bankruptcy predictions, risk assessments of mortgage and other loans

The prediction of corporate bankruptcies is an important and widely studied topic since it can have significant impact on bank lending decisions and profitability. Banks need to predict the possibility of default of a potential counter-party before they extend a loan. They also need to assess the credit risk of bank loan portfolios.

The traditional approach for banks for credit risk assessment is to produce an internal rating, which takes into account various quantitative as well as subjective factors, such as earnings of a person, his/her reputation, etc., via a scoring system. The problem with this approach is of course the subjective aspect of the prediction, that makes it difficult to calculate consistent estimates. Many banks, especially the smaller ones, use the ratings issued by the standard credit rating agencies, such as Moodys and Standard &Poors for credit risk assessment.

Research studies on using ANNs for bankruptcy prediction started in 1990, and are still active now. ANNs provide a nonlinear approach to solve the bankruptcy prediction problem which is much superior to supervised linear prediction methods such as linear regression. Many banks have developed and are using proprietary ANN prediction models instead of traditional approaches.

B. Credit Card Fraud Detection

Various studies and data indicate that many people in the past have become victims of credit card frauds. Now-a-days, internet has become an important part of humans life. A person can buy, invest, and perform many banking tasks online. Almost all the banking organizations have their own website, where customer can perform all the tasks like shopping, etc. All they need to do is provide their credit card details. With progress of technology and ease of use of online banking, many e-commerce organizations have been experiencing the increase in credit card transaction and other modes of online payments. Because of this credit card fraud has become a very popular issue for credit card industry. It causes many financial losses to the customer and also to the organization.

There have been a few studies conducted to determine how to detect frauds by using different technology. One of the popular methods is using artificial neural networks which can integrate an expert's experience into the software and provide support to the banks to detect frauds.

Past history of transactions like card holder's spending pattern from the previous transactions database, and other inputs like location, income, daily living expense etc. can be compared with the current transaction details to detect credit card frauds via ANN. Large deviations on transactions can indicate a fraud.

C. Voice and Speech Recognition

Artificial Neural Network learning paradigm is also applicable to sequential data like voice and gesture. This technology is used by many financial institutions for user identification and other biometric needs.

Speech recognition involves recognizing individual sounds in the audio. To do so sequences of sounds need to match with existing words, and the sequences of words should make sense in some language. This is called language modeling. Language models are trained over large corpora of text using artificial neural networks. These models are finally used to predict voice and speech patterns for a person. Hence like many other complex services, neural networks have found its way in voice and speech recognition.

IBM research has been working on using deep neural networks in Voice and Speech Recognition. [Info. Source: Lecture by Dr. Chung-Sheng Li]

V. ADVANTAGES AND DISADVANTAGES OF ARTIFICIAL NEURAL NETWORKS

Neural networks offer a number of advantages. This includes requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between independent and dependent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms. The unique characteristics of ANNs - non-linearity, adaptability, arbitrary function mapping ability - make them quite suitable and useful for forecasting tasks. Key advantages of ANNs for financial forecasting problems can be listed as:

- Generalization: ANNs dont memorize the training data but they learn the existing patterns, so that they can generalize from the training data set to new data. This is necessary as the data set for any time series forecasting problem keeps varying with time.
- Trainability: ANNs can learn to form associations between the input and output patterns. This is useful as it can used to teach the network to classify stocks, forex etc into correct categories.
- Non-linearity: ANNs can compute non-parametric, nonlinear functions of their input. This enables them to perform complex transformations on data. This is useful since most time series forecasting problems comply to a nonlinear process.
- Uniformity: ANNs offer a computationally uniform paradigm. This enables ANN to easily integrate constraints from various types of inputs. For example, an ANN can easily train on both basic and differential inputs.
- Robustness: ANNs are tolerant of noisy data as well as physical damage. In fact, sometimes noisy data helps ANNs to form better generalizations. This is valuable as a forecasting problem data may include noise at various input levels.
- Parallelism: ANNs are highly parallel in nature. This makes them suitable for implementations on massively parallel computers. This permits super fast processing of both training as well as test data.

A considerable amount of research has been done on ANNs and they are found to give satisfactory performance. However, the findings are inconclusive as to whether and when ANNs are better than classical methods. ANNs do offer a promising alternative approach to traditional linear methods. However, ANNs also embody a large degree of uncertainty. Like statistical models, ANNs have strengths as well as weaknesses. Disadvantages of ANNs include:

- Black box nature of ANNs
- Greater computational burden
- Proneness to overfitting
- Empirical nature of ANN model development

VI. CHALLENGES FOR ARTIFICIAL NEURAL NETWORKS IN FINANCIAL FORECASTING

A. Implementation Challenges

Despite the many satisfactory characteristics of ANNs, building a neural network forecaster for a specific forecasting problem is a nontrivial task. All modelling issues that affect the performance of an ANN must be considered carefully. One critical decision is to determination of appropriate architecture which includes the number of layers in the neural network, the number of nodes in each layer, and the number of arcs, which interconnect the nodes. Other neural network design decisions include but may not be limited to: - the selection of activation functions of the hidden and output nodes, the training algorithm, data transformation or normalization methods, training and test sets, and performance measures.

B. Adoption and Regulatory challenges

Although Neural Networks have proved to be quite promising, the challenge is to decide whether the ANN process is sufficient and defensible in court. By nature, predictive algorithms are often conflict-based, dispute-based or highly sensitive, and the perception that predictive coding introduces uncertainty hinders its adoption. Despite this challenge, recent court cases have shown support on promoting the use of predictive coding. Financial institutions must take the following key points into consideration before using ANNs to solve a problem (INFO source [5]):

1) Assessing how the neural network Algorithm and sampling method can be applied to solve the problem i.e., determine the criteria for measuring the precision of a sampling process and how workflow steps improve the ANN.
2) Learn "how it all works," because any user using this functionality will experience the approach to be quite complex and also fear that the courts will not accept the approach or understand its accuracy.

3) Inform the opposing party beforehand whether and which predictive coding technology will be leveraged for the given problem, and also provide a summary on how it will be used. This is important to do as transparency and communication to opposing counsel are considered very important.
4) Improvement of success rates in applying artificial neural networks and other predictive technology must be achieved by streamlining data collection and preservation of processes.

VII. KEY PLAYERS

Artificial Neural Networks are used by many companies for various applications. Goldman Sachs, Merill Lynch, JP Morgan, KPMG, Chase Manhattan Bank, American Express are some of the companies that efficiently use ANNs in solving their financial and investing problems.

Goldman Sachs [8] and JP Morgan [9] have deployed neural network based trading systems. California Scientific Software has developed a system called BrainMaker[9] that uses neural networks for financial forecasting and trading. Brain Trader Investments [9] of Cambridge, Massachusetts, also employs a proprietary neural networks system called Brain. Merill Lynch is using neural networks for bond-pricing[6] and other time series forecasting issues.

NeuroShell Trader is a pioneer software built by Ward Systems Group. It is based on artificial neural networks and can be used for time series forecasting problems such as building stock market, futures, index and forex trading systems. Some of the worlds most respected software companies using this software include: - Merill Lynch, Citi, Bank of America, Chase, Prudential Securities, New York Life and GE Capital [14].

Another example is of KPMG who has used neural networks to deploy a system for Inherent Risk Analysis(IRA) to analyze the inherent risk of audit clients in manufacturing and merchandising industries [11]. American Express is making use of deep neural network analytics for credit card fraud detection [12].

Another example is of a statistical-based hybrid neural network deployed at Chase Manhattan Bank is one of the largest and most successful AI applications in the United States [10]. It addresses a critical success factor in the bank's strategic planning reducing losses on loans made to public and private corporations. Most of Chase's business for corporations involves assessing their creditworthiness. Chase loans $300 million annually. It has long searched for tools to improve loan assessment. This assessment allows Chase to mitigate risk and seek out new business opportunities.

Internationally many banks such as National Bank of Romania [1] and others use ANN for financial forecasting (exchange rates etc.).

A. *FINRA AND MARKET REGULATION*

[Info. Source: Lecture by Dr. Partha Kanjilal]

FINRA—the Financial Industry Regulatory Authority—is an independent, non-governmental regulator for all securities firms doing business with the public in the United States of America. Its mission is to provide protection to investors and integrity of market.

One such application is Cross Market Equity Surveillance. It does so by conducting automated surveillance using a suite of comprehensive Cross Market equity surveillance patterns that aggregate data across all FINRA-regulated equity markets.

Currently FINRA captures 200k trades as potential manipulative trades. It uses the following machine learning approach to identify potential manipulative trades:

1) Selecting attributes
2) Using current logic as well as unsupervised learning methods such as clustering to label the trade data
3) Ensemble modeling to identify and label suspect trades which should be investigated

Since trading involves buy and sell rates to fluctuate non-linearly with time, ensemble modelling is usually done using supervised learning methods such as Artificial Neural Networks, Decision trees etc.

VIII. OUTLOOK AND FUTURE TRENDS

Artificial Neural Networks are good for a variety of problems, both inside and outside finance for finding trends in large quantities of data. But the question is, can they ever fully simulate the human brain? Can they make a machine conscious of its own existence? May be yes, maybe not. It all depends on how the technology emerges in the future.

As reported in Gartner [20], ANNs or deep neural nets technology seems to have inflated expectations (Fig. 5). The application of ANNs remain largely irrelevant to 95% of current use cases both inside and outside financial industry.

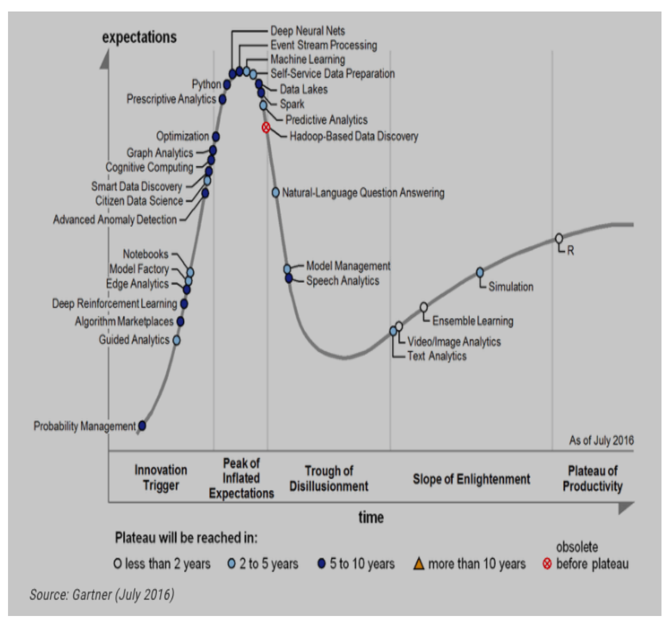

Fig. 5. Hype Cycle for Data Science, 2016, Image Source[20]

However, outlook for ANNs remains very positive and many companies have already started piloting them. The results will be much more clearly visible 5 years down the line. Let us discuss some of the advances and future trends in artificial neural networks:

1) Integrating Fuzzy Logic in artificial neural networks: Fuzzy logic refers to a type of logic that recognizes more than simple true and false values. For example, today is a sunny day is 100% true if there are no clouds, 80% true if there are some clouds, 50% true if its hazy and 0% true if its raining and no sun is visible. Introducing this logic into stock price prediction can help us precisely predict the amount by which the price of stock may rise or fall.
2) Hardware specialized for neural networks: Many semiconductor design groups and organizations, engineering and design firms, system research, and military and intelligence organizations are researching on neuromorphic and competing hardware designs that could drastically improve the scope and run time of artificial neural networks [13].
3) Pulse Neural Networks: Biological Neural Networks communicate through pulses and use the timing of the pulse to perform computation and communicate information. This realization by researchers has stimulated significant research on pulsed neural networks opening an unexplored domain for neural networks.

IX. Conclusion

In this paper we described the theory behind artificial neural networks and its salient features. We also discussed how ANNs are used for time series forecasting problems, advantages, disadvantages and challenges of using ANNs, key players using this technology and finally the expected future trends in ANNs.

Artificial Neural Networks are quite promising and will continue to provide disruptive solutions over the coming years. The computational power of artificial neural networks, their ability to handle almost endless amounts of data, and unprecedented advances will allow financial institutions to harness data in order to adapt to new problems and situations that no one has encountered previously. Thus ANNs will continue to be an effective prediction tool over years to come.

References

[1] Bogdan Oancea, Stefan Cristian Ciucu, Time Series Forecasting using Neural Networks, arXiv preprint arXiv:1401.1333, 2014

[2] Wanjawa, Barack Wamkaya , Muchemi, Lawrence, ANN Model to Predict Stock Prices at Stock Exchange Markets, arXiv preprint arXiv:1502.06434, 2014

[3] Sanjay Singh, Anchal Karnwal, Neha Prasad, Royston DSouza, Ashwin Shenoy, Fraud Detection using Neural Network, Conference Paper, Oct 2014

[4] Jack V. Tu, Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes, Journal of Clinical Epidemiology, Volume 49, Issue 11, November 1996, Pages 1225-1231

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

[6] http://www.bloomberg.com/news/articles/1992-11-01/where-neural-networks-are-already-at-work

[7] Dimitrios Maditinos; Prodromos Chatzoglou, The use of Neural Networks in Forecasting, Review of Economic Sciences, 6, TEI of Epirus, pp. 161-176

[8] Stephen Prentice: Defining Algorithmic Business, Gartner Report, 2016

[9] http://www.financialexpress.com/archive/neural-networks-make-money/204693/

[10] http://ulcar.uml.edu/ iag/CS/intronn.htm

[11] Allen Kent, James G. Williams, Encyclopedia of Microcomputers: Volume 9

[12] https://www.capgemini.com/resourcefileaccess/resource/pdf/Credit_Card_Transaction_Fraud_and_Mitigation_Trends.pdf

[13] Mike J. Walker, Betsy Burton, Michele Cantara, Hype Cycle for Emerging Technologies, Gartner Report, 2016

[14] http://try.neuroshell.com/index/?gclid=CNr3kvuEotACFYlWDQodlLcECQ

[15] https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Future/index.html

[16] Shital Solanki, H.B.Jethva, A review on back propagation algorithms for Feedforward Networks, Volume : 2, Issue : 1, Jan 2013 ISSN No 2277 - 8160

[17] FernandoJ, Pineda, Generalization of Backpropogation to recurrent and higher order neural networks, American Institute of Phvsics 1988

[18] Jan A, Snyman (2005), Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms, Springer Publishing, ISBN 0-387-24348-8

[19] Abhishek Kar, Stock Prediction using Artificial Neural Networks, Dept. of Computer Science and Engineering, IIT Kanpur.

[20] Peter Krensky, Alexander Linden, Jim Hare, Hype Cycle for Data Science, Gartner Report, 2016