**2. Literature Review**

This chapter presents the literature and previous activities that are most relevant for this work project. The literature review is done accordantly with the most relevant subjects that were taking into account for the preparation of this project, more specifically Machine Learning techniques to forecast stock prices.

In the last few decades forecasting of stock returns has become an important field of research. In most of the cases the researchers had attempted to establish a linear relationship between the input macroeconomic variables and the stock returns.

After the discovery of nonlinearity in the stock market index returns, many literatures have come up in nonlinear statistical modeling of the stock returns, most of them required that the nonlinear model be specified before the estimation is done. Youngohc Yoon and George Swales in1991 demonstrated that the neural network approach is capable of learning a function that maps inputs to output and encoding it in magnitudes of the weights in the network connection. And compared Neural Network technique with Application of Artificial Neural Network for stock market predictions: A review of literature International Journal of Machine Intelligence, Multivariate Discriminate Analysis approach Indicated that the Neural Network approach can significantly improve the predictability of stock price performance.

ANN has evolved out to be better technique in capturing the structural relationship between a stock’s performance and its determinant factors more accurately than many other statistical techniques. In literature, different sets of input variables are used to predict stock returns. In fact, different input variables are used to predict the same set of stock return data. . Some researchers even preprocessed these input data sets before feeding it to the ANN for forecasting.

**2.1 Relevant Works**

Wilson and Sharda studied prediction firm bankruptcy using neural networks and classical multiple discriminant analysis, where neural networks performed significantly better than multiple discriminant analysis.

Min and Lee were doing prediction of bankruptcy using machine learning. They evaluated methods based on Support Vector Machine, multiple discriminant analysis, logistic regression analysis, and three-layer fully connected back-propagation neural networks. Their results indicated that support vector machines outperformed other approaches. Lee was trying to predict credit rating of a company using support vector machines. They used various financial indicator and ratios such as interest coverage ratio, ordinary income to total assets, Net income to stakeholders’ equity, current liabilities ratio, etc. and achieved accuracy of around 60%. Predicting credit rating of the companies were also studied using neural networks achieving accuracy between 75% and 80% for the United States and Taiwan markets.

Tsai and Wang did a research where they tried to predict stock prices by using ensemble learning, composed of decision trees and artificial neural networks. They created dataset from Taiwanese stock market data, taking into account fundamental indexes, technical indexes, and macroeconomic indexes. The performance of Decision Tree + Artificial Neural Network trained on Taiwan stock exchange data showed F-score performance of 77%. Single algorithms showed F-score performance up to 67%.

Kim and Han used a genetic algorithm to transform continuous input values into discrete ones. The genetic algorithm was used to reduce the complexity of the feature space. This paper proposes a novel evolutionary computing method called a genetic quantum algorithm. Genetic Quantum Algorithm is based on the concept and principles of quantum computing such as qubits and superposition of states. Instead of binary, numeric, or symbolic representation, by adopting bit chromosome as a representation Genetic Quantum Algorithm can represent a linear superposition of solutions due to its probabilistic representation. As genetic operators, quantum gates are employed for the search of the best solution.

The similar work in cash forecasting of a bank branch was implemented in MATLAB by Premchand and Ekta [2006]. Neural networks are used to analyze the system. The system performs better than other systems based on time series. Its performance was also better than one of the available Excel Add-in for forecasting “Alyuda Forecaster XL 2.3”. This system can be scaled for all branches of a bank in an area by incorporating historical data from these branches. Such a system will help the bank for proper and efficient cash management.

Survey of existing literature reveals that there are different types of ANN models used for predicting the stock market. Many researchers noted that slight parameter changed causes major variations in the behaviour of the network. So there is no theory which could be guideline for finding best network topology. Recently, Pratap and Ambika [2011] proposed trigonometric functional link artificial neural network (FLANN) model employs standard least mean square (LMS) algorithm with search-then-converge scheduling. The network could effectively calculate learning rate parameter that changes with time and may require less experiments to train the model. Here FLANN is used for long term as well as short term stock market prediction

There are many tools and software available out there that provide forecasting of stock market entities. Some of the popular tools and software with their methodologies are mentioned as follows:

**Markettrak**

Its stock market forecast system consists of two major parts: an extensive database and a forecast model. The forecast model reads the database and then makes a prediction of where the market is headed. From this prediction, it determines a trading position for the Dow Diamonds or the SP500 Spiders. The database and forecast are updated daily at the close of trading.

It uses a neural network model in combination with a genetic algorithm to calculate the SP500 forecast. The calculations are somewhat complex but can be summarized by the following three procedural steps.

Step one: The genetic algorithm is used to find the optimum neural network structures and inputs. This calculation basically determines how the networks will be wired.

Step two: Using the information from the first step, a set of networks is initialized and then trained on about 75 percent of the market data (in-sample) in their database, which currently consists of about 7200 days of data. They use an evolutionary program to train the networks (i.e. to determine the neural network weights).

Step three: After training, the networks are rigorously tested on the remaining 25 percent of market data (out-of-sample). Networks that fail the test are discarded. Networks that pass the test are included in the library that they use to calculate the forecast. The number of neural networks currently in their library varies from day to day, but normally contains more than 400.

Inputs to the networks are technical and fundamental market data. The table below shows the types of data that are currently used by the model:

* Dow Jones Industrial Average closing value
* Dow Jones Industrial Average theoretical high value
* Dow Jones Industrial Average theoretical low value
* Dow Jones Transportation Average closing value
* Dow Jones Utility Average closing value
* New York Stock Exchange total volume
* New York Stock Exchange number of advancing stocks
* New York Stock Exchange number of declining stocks
* New York Stock Exchange number of new highs
* New York Stock Exchange number of new lows
* New York Stock Exchange advancing volume
* New York Stock Exchange declining volume
* SP500 closing value
* SP500 trailing earnings
* Yen-Dollar exchange rate
* Treasury bill discount rate
* Commodity Research Bureau index

The above data are filtered and normalized and certain functions of these data are computed. It currently computes 63 separate input variables at the close of each trading day. The 63 inputs are applied to a neural network and after some number crunching the network outputs a value between -1.0 and +1.0, with -1.0 being a very strong down market signal and +1.0 being a very strong up market signal. A value near zero would indicate a neutral market signal. They apply the inputs to each network in our library and an average of their outputs is computed. This average network output is used with position set points to determine a trading position for the Dow Diamonds or the SP500 Spiders for the next trading day. When the computed value of the average network output is above the long position set point, a long position is indicated. When the value of the average network output is below the short position set point, a short position is indicated. When the average network output falls between these two set points, a cash position is indicated. Because of the timing of the update, trades may be made in the extended sessions or at open of the next trading day.

When computing our performance, all trades are assumed to take place at the session close. The current trading position along with recent average network output values and the average network output set points are shown on our forecast page.

**Stock-Forecasting.com**

www.stock-forecasting.com (Center of Mathematics & Science, Inc., Chicago, United States of America) provides innovative price-prediction technology for active Day Traders, Short- and Long-term Investors. They develop web-based software for stock market forecasting and analysis.

The artificial intelligence www.stock-forecasting.com software is based on neural network technology, advanced statistical methods and non-periodic stock price wave analysis. The Stock-Forecasting software predicts stock prices, generates trading "Buy-Hold-Sell" signals, computes the most profitable company to invest in and analyzes the accuracy of predictions.

REFERENCES:

<http://www.markettrak.com/markettrak-bin/FinSystem/forecast.cgi>

<http://stock-forecasting.com/>

<https://run.unl.pt/bitstream/10362/21452/1/TEGI0383.pdf>

<https://www.researchgate.net/publication/320449394_Stock_Market_Prediction_Performance_of_Neural_Networks_A_Literature_Review>