

# Stock Price Prediction and Portfolio Selection Using Artificial Intelligence

Sandeep Patalay<sup>a,\*</sup>, Madhusudhan Rao Bandlamudi<sup>b</sup>

<sup>a</sup> *Research Scholar, Dept. of Management Studies, Vignan's University, India*

<sup>b</sup> *Professor, Dept. of Management Studies, Vignan's University, India*

---

## ABSTRACT

Stock markets are popular investment avenues to people who plan to receive premium returns compared to other financial instruments, but they are highly volatile and risky due to the complex financial dynamics and poor understanding of the market forces involved in the price determination. A system that can forecast, predict the stock prices and automatically create a portfolio of top performing stocks is of great value to individual investors who do not have sufficient knowledge to understand the complex dynamics involved in evaluating and predicting stock prices. In this paper the authors propose a Stock prediction, Portfolio Generation and Selection model based on Machine learning algorithms, Artificial neural networks (ANNs) are used for stock price prediction, Mathematical and Statistical techniques are used for Portfolio generation and Un-Supervised Machine learning based on K-Means Clustering algorithms are used for Portfolio Evaluation and Selection which take in to account the Portfolio Return and Risk in to consideration. The model presented here is limited to predicting stock prices on a long term basis as the inputs to the model are based on fundamental attributes and intrinsic value of the stock. The results of this study are quite encouraging as the stock prediction models are able predict stock prices at least a financial quarter in advance with an accuracy of around 90 percent and the portfolio selection classifiers are giving returns in excess of average market returns.

*Keywords:* Stock Prediction, Stock Portfolio, Artificial Intelligence (AI), Machine Learning (ML), Artificial Neural Networks (ANN), K-Means, Multi-Layer Perceptron (MLP)

---

## I . Introduction

The stocks markets have become an important investment activity around the world and the proliferation of internet and online trading systems have drawn millions of small individual investors in to

the stock market (Tsai and Wang, 2009). For individual investors who are interested to invest their hard earned money in stock markets is major financial decision to earn superior returns compared to other financial instruments. Investment in stock markets needs sound understanding of the risks involved in

---

\*Corresponding Author. E-mail: [sandeep.patalay@gmail.com](mailto:sandeep.patalay@gmail.com)

the stock price evaluation which depends on the technical and fundamental features of the stock. Majority of the investors do not understand the complex dynamics of stock markets nor are they qualified to select stocks for investments. For many small time individual investors it is beyond their budget and time to take the help of good financial advisors who can guide them in stock market investments. The reputation of these financial advisors has also been in question many times due to the wrong decisions they have taken without any consideration for Risk vs Reward calculations. Many of the financial advisory websites related to stock price prediction have not been accurate and have been a cause for financial doom for many small time investors (Leong and Zaki, 2018), (Montmarquette and Viennot-briot, 2012). There is strong need for an Automated Decision Support System which can help individual investors to accurately predict stock prices and automatically create a good portfolio of stocks to diversify risk and earn premium returns.

The Main Objectives of this study are:

- Study the computational systems available for stock price prediction
- Study the current Machine learning techniques used for stock prediction and describe their limitations
- Study the critical intrinsic and macro-economic features that directly affects the stock price
- Develop a Stock prediction Model based on ANNs
- Develop a Stock portfolio model using Machine learning clustering techniques
- Design a Computational Decision Support system to predict stock prices and create automatic stock portfolios
- Describe the outcomes, implications and limi-

tations of the study

Detailed literature review in the area of stock price prediction and portfolio creation finds that there are very few machine learning models based on fundamental attributes of a stock (Huang et al., 2019) and majority of the portfolio selection models still use traditional quantitative models (Dipietro, 2019) and do not use machine learning techniques. The model presented in this paper adds significant body of knowledge to the existing literature by using the machine learning models such as ANNs and K-means using fundamental attributes for stock price prediction and portfolio selection respectively, the literature review section provides the details of the findings.

In this study, we propose a financial decision support system which can predict stock prices with a reasonable degree of accuracy for long term investments as opposed to daily price movements and also perform the risk analysis to automatically create a portfolio of top performing stocks for risk diversification and earning premium returns compared to average market return. The stock prediction is based on special case of Artificial Neural Networks (ANNs) called Multi-Layer Perceptrons (MLP), the MLP model takes in to account the fundamental features which are intrinsic to the stock. The fundamental features also called as Intrinsic features of the stock such consists of critical financial ratios of the firm that the stock represents. The MLP in this study was trained using open source financial data of select stocks listed on the NSE (National Stock Exchange, India). The financial decision support system also features an automatic portfolio creation based on the top performing stock predictions and performs an evaluation of the Risk vs Reward criteria. The automatic portfolio creation is an un-supervised clas-

sification problem and we propose to use K-Means clustering algorithm to select group of stocks based on the Portfolio Risk ( $\beta$ ) and Portfolio return ( $R_p$ ). The study is divided in to two distinct phases; the first phase involves the design of MLP based stock prediction model and the second phase involves creation of stock portfolio based on supervised machine learning classification models. The details of these two phases are described in the “Models and Methods” Section 3 of this paper.

The data for training these models is collected mainly from open source financial websites, The top 50 companies by Market capitalization listed on the NSE were considered for analysis, The data was partitioned in to training and test sets for proper validation. The results of the ANN stock price prediction are accurate and relative root mean squared error of 8.77% and based on the stock price, the portfolio selection was done by an unsupervised learning K-Means clustering algorithm that created 6 portfolios and cluster 1 was selected because of the highest return and the lowest risk profile, details of the data sources and analysis is provided in the “Data Analysis” Section 4 of this paper. The major implications of this study are that ANNs with fundamental stock attributes are more efficient and accurate in predicting stock prices and using K-means clustering algorithm for stock portfolio selection provides premium returns and reduces the portfolio risk. This study has implications on the financial domain where machine learning models in financial decision support can make the capital markets more efficient and less risky.

## II. Literature Review

The main aim of this literature review is to find

out prior research in the area of stock prediction and particularly find out the methods and models used in financial domain for stock price prediction. The research papers are reviewed based on the Machine Learning or Mathematical model they employ, Input feature set used to train models based on technical or fundamental attributes, Model Outputs in the form of classification signals (Trend, Buy, Sell, Hold, Etc.) or Numerical signals (Stock Prices) and prediction time frame (i.e., Long term or Short term prediction).

The basic premise that Stock prediction is a viable option can be derived from the outcome of studies involving Efficient Market Hypothesis (EMH). As per the EMH theory (Malkiel, 2003), financial markets are considered to be very efficient and the asset prices reflect all the financial information available in public domain and it is impossible to generate excessive returns through informed investment decisions. The EMH theory believes that market participants make rational decisions and any deviations will be quickly eliminated by market arbitrageurs.

(Sung et al., 2016) make significant contribution to prove that Efficient Market Hypothesis is null and proves that the machine learning techniques such as ANNs and SVMs are able to forecast financial markets. The study compares the legacy econometric Models adopted by Economists vs cutting edge prediction models based on Machine learning. This is a foremost study conducted on stock markets to compare effectiveness of ML vs Econometric prediction models on short term Daily and Intraday stock prices. The results of this study have major implications in the field of Finance as they suggest that the Machine learning models like ANNs and SVMs outperform the econometric models which are mostly linear in nature. The ML techniques can build complex non-linear models which are suitable for stock

predictions.

A detailed survey of literature was done on current stock prediction computational methods and usage of Decision Support systems in financial domain. There are few important methods for stock prediction namely (1) Technical Analysis, (2) Time Series Forecasting, (3) modeling and predicting volatility of stocks using differential equations and (4) Machine Learning and Data Mining (Khaidem et al., 2016). This literature review focuses on the Machine Learning and Data Mining techniques.

In the last decade, there has been a focused effort both in Academia and Industry to explore stock price trend prediction from finance perspective as well as from the combination of two major IT areas which are AI and Data Mining (Yong and Taib, 2009). In the recent times, with the proliferation of Machine learning algorithms in the financial domain, many studies have focused on the usage of various machine learning algorithms including ANNs for stock price prediction, however there are many limitations in the current application of ANNs which have mostly focused on the technical parameters which result in inaccurate and unusable prediction by investors (Pyo et al., 2017). Many studies in this field have suggested the use of various computational models for stock price prediction such as ANNs, SVMs and Genetic algorithms which have mostly used technical indicators (Ican and Çelik, 2017). Recent studies on the comparison of various computational and machine learning models used for prediction of stock market indices have indicated that ANN's performance is better than other computational models (Banik et al., 2014).

The price of stock price involves complex combination of company performance, investor sentiments and economic factors. (Kusuma et al., 2019) have developed a model using ANNs with candlestick tech-

nical indicators as input and predicting the future movements of the stock price.

(Pang et al., 2018) have developed a method using deep long short-term memory neural network (LSTM) and long short-term memory neural network with automatic encoder to the stock market index. The input features are purely based on technical indicators. The accuracy of two models is poor with 57.2% and 56.9%, respectively, for the Shanghai A-shares composite index. This method is not suitable for long term invest decisions as the predictions are only useful for day Traders.

(Di Persio and Honchar, 2016) propose a machine learning approach to predict stock market indices using various ANN techniques such as MLP, CNN, Wavelet NN and other ensemble techniques. The main goal was to predict the stock market index trend based on technical parameters and features the time series analysis and forecasting. The accuracy of the model is quite poor with only around 50% accuracy. Based on the results, it claims that ANNs are most suitable in dealing with financial data. This approach is only suitable for Short term and Day traders.

(Chong et al., 2017) propose a Stock Prediction Model based on Deep Neural Networks (DNN) with intraday technical features as input to the model and is focused on forecasting the times series stock price data with the help of Technical indicators. The study employs three data representation methods; principal component analysis, auto encoder, and restricted Boltzmann machine, and construct three-layer deep neural networks (DNN). The main finding of this study is that Artificial Neural Networks (ANN) performs better than the linear regression models.

(Malagrino et al., 2018) explore the options of predicting Stock Market index movements using a Machine learning technique called Bayesian Networks.

The study focusses on the external events that affect the Brazilian stock market index (iBovespa) such as closing directions of other stock market indices. The accuracy of the Bayesian network is stated as 71%. The prediction of stock market indices does not give any benefits to long term investors in terms of investment decision making process.

Some of the recent systems developed for financial domain prediction are more focused on the broad Buy, Sell or Hold signals as opposed to numerical prediction of stock prices, these methods are simpler than the numerical prediction models and are widely being used as stock trading systems, though their accuracy levels are low. These systems are based on Machine Learning techniques such as Naive Bayes, k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Artificial Neural Network (ANN) and Random Forest (Selvamuthu et al., 2019). Time series technical data is used as input features for the Machine learning models. A separate model for each stock is trained on the past time series technical data. A major shortfall of this approach is that it does not generate sell signals when the stock prices have gone down (Report, 2016). Using a similar methodology (Khaidem et al., 2016) give a brief summary of the various stock prediction techniques such as Technical analysis, Time series forecasting, Machine Learning and differential equations. The authors propose a novel approach to predict stock prices based on a machine learning technique called as Ensemble learning which is in turn based on Random Forest decision trees. The input feature set to the model is technical data such as Relative Strength Index (RSI), stochastic oscillator, Etc. (Agrawal et al., 2019) have developed a stock prediction model using ANNs that outputs the broader trend of BUY, SELL or HOLD signals and conclude that ANNs are suitable for financial domain.

There are also studies which use hybrid machine learning techniques to boost the performance of stock market predictions. (Tsai and Wang, 2009) propose a novel methods of stock price prediction by combining two machine learning methods namely the Decision trees and ANNs. The combined model seems to be more accurate than single models with accuracy in the range of 70%. The inputs feature set for the model consists of technical time series data.

There has been some interesting work in the area of using daily stock trading patterns to predict stock prices. Detecting trading patterns in a stock market and forecasting the stock movements is a complex task given the non-linear and non-stationary behavior of the stock prices. Stock markets generate a complex multi frequency trading patterns which eventually affect the stock prices. (Zhang et al., 2017) have proposed a novel method of combining machine learning techniques with core mathematical principles. The Neural networks are used in combination of Fourier transforms to create a State Frequency Memory (SFM) recurrent network used for stock price prediction (Zhang et al., 2017). The accuracy of this system is still unproven given the complexities involved in building such a model (Wen et al., 2019).

There are many limitations in the current machine learning computational models including ANNs used for stock price prediction (Diakoulakis et al., 2018), the inaccuracies are mainly because most of these models do not take in to account the fundamental parameters affecting the stock price as they are mostly rely on daily technical indicators which are very volatile. Anbalagan & Maheswari show that the Simple Moving Average (SMA), Exponential Moving Average (EMA), Moving Average Convergence Divergence (MACD) and Relative Strength Index (RSI) are some of the Technical Indicators which

are used as input to train the ANN system which is integrated with Fuzzy Metagraph (Anbalagan and Maheswari, 2014).

In the existing literature on financial domain, there are very few stock prediction models which are based on Fundamental attributes of the stock. One such paper (Cheng and Chen, 2007) by provides a methodology to use stock fundamental features to predict Revenue growth rate (RGR) which is one of the most important factor in determining stock price. The prediction models are developed using Decision Trees, MLP, Bayes Networks and Rough sets. The input features to model are revenues, assets, income, profit and costs. This model gives a good baseline to use fundamental analyses for stock price prediction. (Huang et al., 2019) have developed a comparative study of feed-forward neural network (FNN) and adaptive neural fuzzy inference system (ANFIS) based on fundamental attributes and conclude that FNNs have better performance over other techniques.

Many ANN models have used technical indicators as the input feature set and have presented detailed limitations with the models (Dase and Pawar, 2010). The studies which have used ANNs for stock price prediction with fundamental parameters have limited themselves to small subset of features such as EPS (Earnings per Share) (Rudin, 2012).

An empirical study conducted show that the selection of input variables can generate higher forecast accuracy and it is possible to enhance the performance of the optimized ANN model by selecting input variables appropriately (Qiu and Song, 2016). Many studies have analyzed existing and new methods of stock market prediction and have identified a common flaw in Technical Analysis methodology (Dunne, 2017). It is very important to select the input parameters for ANNs which have a direct effect on the

stock price which depend on the intrinsic financial ratios of the stock.

On the other hand, Stock portfolio generation, selection and Optimization still use traditional quantitative techniques (Dipietro, 2019) and a lesser usage of Machine learning, though large fund managers in developed economies have started using these techniques with limited results (Song, 2014). (Flechas Chaparro et al., 2019) have done comprehensive studies of the quantitative and econometric models that are widely used in financial domain for stock portfolio selection and their evolution, it can noted that quantitative models are still widely used in for stock portfolio selection. The race is on to use Machine learning techniques for Stock Portfolios (Text, 2018), but many of these initiatives are limited to predicting the general trend and the optimization of portfolios (Arik et al., 2014) rather than constructing a complete portfolio which can perform a completely automatic Risk and Reward analysis.

After doing extensive literature review, we find that there are many computational and machine learning models used in stock price prediction, many of these models have focused on determining the daily movements of stock prices and mostly rely on technical indicators and ignore the fundamentals driving the stock price. Most of the models that we have reviewed in the literature have only focused on day-to-day price prediction which is of interest to Day traders, but we could not find any comprehensive models that were built for value investors who are more focused on long term investments. Based on this finding, we believe that there are very few comprehensive machine models currently to predict long-term movements of stock prices which shall be very useful for individual stock investors. Also the current Stock Portfolio construction, selection and optimization models are mostly manual and the

use of machine learning in this area is yielding limited results.

Based on the literature survey, ANNs are the best approach for stock price prediction after extensive comparative analysis between various computational models and find that choosing the correct input feature set to the ANN consisting of fundamental financial parameters can improve the accuracy of stock price prediction and we conclude that there is a strong need for developing a Machine learning model based on ANNs with intrinsic features as input set and also generate and select stock portfolios with Risk and Reward analysis done by machine learning algorithms.

### III. Methods and Models

In this study, Machine learning and Mathematical Models are used for Stock price Prediction, Portfolio Generation, Evaluation and Selection. The Model consists of three distinct phases where the first phases involves predicting the Stock Prices using Machine

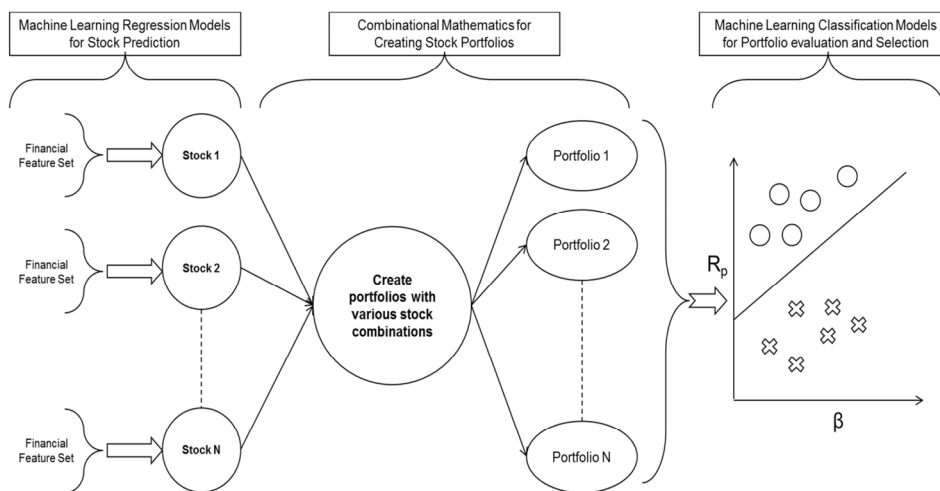
Learning ANNs, Second phase involves auto generation of Portfolios consisting of various stock combinations which is based on a Mathematical Model and the final phase involves using supervised Machine learning Classification Models to evaluate the Portfolios and finally select the best performing Portfolios. The Architecture of the system for predicting stock prices, automatic stock portfolio generation, evaluation and selection is described below. The main components of the system consist of:

- ANN based Stock prediction Model
- Mathematical Model for creating Stock Portfolios
- Clustering Model for Stock Portfolio Evaluation and Selection

#### 3.1. ANN Based Stock Prediction Model

##### *Fundamental Value / Intrinsic Value of a Stock*

Intrinsic value refers to the value of a stock determined through fundamental analysis without reference to its market value (Putra et al., 2019). It is also generally called fundamental value. The attrib-



<Figure 1> Machine Learning Based Stock Price Prediction and Portfolio Selection Model

utes that directly affect the intrinsic value of a stock are its core financial indicators that indicate the health of the company and its long term growth prospects. The financial parameters groups considered for intrinsic value are listed below:

- Investment Valuation Ratios ( $R_v$ )
- Profitability Indicator Ratios ( $R_{pf}$ )
- Liquidity Measurement Ratios ( $R_l$ )
- Debt Ratios ( $R_{db}$ )
- Cash Flow Indicator Ratios ( $R_{cf}$ )

The above ratios represent the Profitability, Growth prospects, financial stability and viability, cash flows and comparative fair value ratios of stocks.

**Investment Valuation Ratios ( $R_v$ ):** Investment Valuation Ratios looks at a wide array of ratios that can be used by investors to estimate the attractiveness of a potential or existing investment and get an idea of its valuation.

- Per Share Data (Basic EPS)
- Price/Book Value Ratio (P/B)
- Price/Cash Flow Ratio (P/CF)
- Price/Earnings Ratio (P/E)
- Price/Earnings To Growth Ratio (PEG)
- Price/Sales Ratio (P/S)
- Enterprise Value Multiple (EV/EBITDA)

**Profitability Indicator Ratios ( $R_{pf}$ ):** Profitability is a key piece of information that should be analyzed when you're considering investing in a company. This is because high revenues alone don't necessarily translate into dividends for investors (or increased stock prices, for that matter) unless a company is able to clear all of its expenses and costs. Profitability

ratios are used to give us an idea of how likely it is that a company will turn a profit, as well as how that profit relates to other important information about the company.

- Return On Assets
- Return On Equity
- Return On Capital Employed
- Earnings Yield
- Dividend Yield

**Liquidity Measurement Ratios ( $R_l$ ):** Liquidity is a measure of how quickly a company's assets can be converted to cash. Liquidity ratios can give investors an idea of how capable a company will be at raising cash to purchase additional assets or to repay creditors quickly, either in an emergency situation, or in the course of normal business.

- Current Ratio
- Quick Ratio
- Cash Ratio
- Cash Conversion Cycle

**Debt Ratios ( $R_{db}$ ):** Debt ratios give investors a general idea of the company's overall debt load as well as its mix of equity and debt. Debt ratios can be used to determine the overall level of financial risk a company and its shareholders face. In general, the greater the amount of debt held by a company the greater the financial risk of bankruptcy.

- Debt Ratio
- Debt-Equity Ratio (D/E)
- Capitalization Ratio
- Interest Coverage Ratio
- Cash Flow To Debt Ratio (CF/D)



**Cash Flow Indicator Ratios ( $R_{cf}$ ):** Cash flow indicators focus on the cash being generated in terms of how much is being generated and the safety net that it provides to the company. These ratios use cash flow compared to other company metrics to determine how much cash they are generating from their sales, the amount of cash they are generating free and clear, and the amount of cash they have to cover obligations

- Operating Cash Flow/Sales Ratio
- Free Cash Flow/Operating Cash Ratio
- Cash Flow Coverage Ratio
- Dividend Payout Ratio

The mathematical function that represents the Intrinsic Value of a Share ( $I_v$ ) is given below:

$$I_v = \varphi (R_v, R_{cf}, R_l, R_{pf}, R_{db}) \quad (1)$$

The above ratios are passed through the attribute selector machine learning algorithm described in section 4.1 “Data Analysis” of this paper, the selection attributes are further formed as input to the ANN for training the stock prediction model.

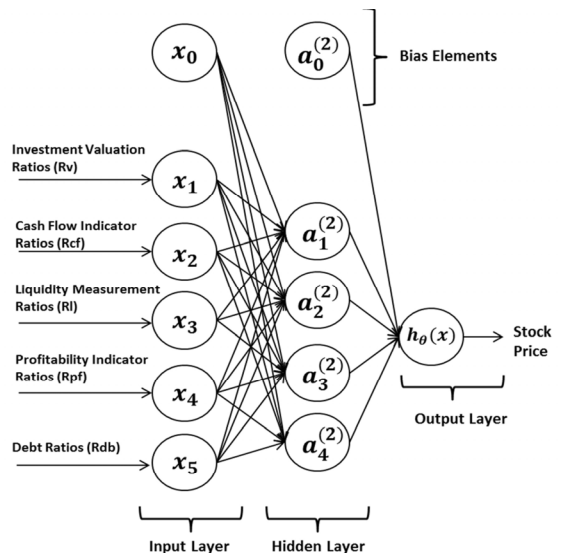
### ANN Model

Artificial neural networks (ANNs) or connectionist systems are computing systems vaguely inspired by the biological neural networks that constitute animal brains. Such systems “learn” tasks by considering historical data, generally without task-specific programming. An artificial neural network is an interconnected group of nodes, akin to the vast network of neurons in a brain. In <Figure 2>, each circular node represents an artificial neuron and an arrow represents a connection from the output

of one artificial neuron to the input of another. Each connection between artificial neurons can transmit a signal from one to another. The artificial neuron that receives the signal can process it and then signal artificial neurons connected to it.

The ANN model used for this study is the three layered Multilayer Perceptron consisting of one input layer, one hidden layer and one output layer. The Multilayer perceptron model is used to benefit from the non-linear features of the hidden layer. The output layer consists of a transfer function and predicts the stock price based on the weights assigned to the hidden layer neurons.

<Figure 2> shows a network with three layers, the first layer of the network represents the attributes in the data. The input layer has an additional constant input called the bias. The 2nd layer is called “hidden” because the units have no direct connection to the environment. This layer is what enables the system to represent share prediction model.



<Figure 2> ANN Model for Share Price Prediction

$$a_1^{(2)} = g(\theta_{10}^{(1)}x_0 + \theta_{11}^{(1)}x_1 + \theta_{12}^{(1)}x_2 + \theta_{13}^{(1)}x_3 + \theta_{14}^{(1)}x_4 + \theta_{15}^{(1)}x_5) \quad (2)$$

$$a_2^{(2)} = g(\theta_{20}^{(1)}x_0 + \theta_{21}^{(1)}x_1 + \theta_{22}^{(1)}x_2 + \theta_{23}^{(1)}x_3 + \theta_{24}^{(1)}x_4 + \theta_{25}^{(1)}x_5) \quad (3)$$

$$a_3^{(2)} = g(\theta_{30}^{(1)}x_0 + \theta_{31}^{(1)}x_1 + \theta_{32}^{(1)}x_2 + \theta_{33}^{(1)}x_3 + \theta_{34}^{(1)}x_4 + \theta_{35}^{(1)}x_5) \quad (4)$$

$$a_4^{(2)} = g(\theta_{40}^{(1)}x_0 + \theta_{41}^{(1)}x_1 + \theta_{42}^{(1)}x_2 + \theta_{43}^{(1)}x_3 + \theta_{44}^{(1)}x_4 + \theta_{45}^{(1)}x_5) \quad (5)$$

$$h_\theta(x) = \theta_{10}^{(2)}a_0^{(2)} + \theta_{11}^{(2)}a_1^{(2)} + \theta_{12}^{(2)}a_2^{(2)} + \theta_{13}^{(2)}a_3^{(2)} + \theta_{14}^{(2)}a_4^{(2)} \quad (6)$$

Where,

$a_i^{(j)}$  Represents the activation unit “i” in jth layer and

$\theta_i^{(j)}$  Represents the weights between interconnecting nodes

$h_\theta(x)$  Represents the output function that represents the share price  $Pe(S)$

The ANN is trained using by adjusting the weights of the connections between each neuron nodes by using the backpropagation algorithm, the solution is to modify the weights of the connections leading to the hidden units based on the strength of each unit’s contribution to the final prediction. The hidden layer uses the “Sigmoid” transfer function which gives the nonlinear learning advantage.

A cost function is a measure of “how good” a neural network did with respect to its given training sample and the expected output. It also may depend on variables such as weights and biases. The cost function equation is given below:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2 \quad (7)$$

Where,

$J(\theta)$  is the cost function that represents the difference between the predicted value and the actual value

$h_\theta(x)$  Represents the output function that represents the share price

$y^{(i)}$  Represents the share price from the training set

“m” Represents the number of training examples

Backpropagation is a method used in artificial neural networks to calculate a gradient that is needed in the calculation of the weights to be used in the network. It is commonly used to train deep neural networks, a term referring to neural networks with more than one hidden layer. The “backwards” part of the name stems from the fact that calculation of the gradient proceeds backwards through the network, with the gradient of the final layer of weights being calculated first and the gradient of the first layer of weights being calculated last. Partial computations of the gradient from one layer are reused in the computation of the gradient for the previous layer. This backwards flow of the error information allows for efficient computation of the gradient at each layer versus the naive approach of calculating the gradient of each layer separately.

$$\text{Given Training Set } \{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\} \quad (8)$$

Where,

$x^{(i)}$  Represents the financial ratios of the i<sup>th</sup> training example

$y^{(i)}$  Represents the share price of the i<sup>th</sup> training example

$$\text{Gradient Computation: } \min_{\theta} j(\theta) = \frac{\partial}{\partial \theta_{ij}^{(l)}} J(\theta) \quad (9)$$

$$\text{Set } \Delta_{ij}^{(l)} = 0 \text{ (for all } l, i, j)$$

$$\delta^{(L)} = \mathbf{a}^{(L)} - \mathbf{y}^{(i)} \tag{10}$$

Compute  $\delta^{(L-1)}, \delta^{(L-2)}, \dots, \delta^{(2)}$

$\delta^{(L)}$  = error of node  $j$  in layer  $l$

$$\Delta_{ij}^{(l)} = \Delta_{ij}^{(l)} + \mathbf{a}_j^{(l)} \delta_i^{(l+1)} \tag{11}$$

$$D_{ij}^{(l)} = \frac{1}{m} \Delta_{ij}^{(l)} + \lambda \theta_{ij}^{(l)} \text{ if } j \neq 0 \tag{12}$$

$$D_{ij}^{(l)} = \frac{1}{m} \Delta_{ij}^{(l)} \text{ if } j = 0 \tag{13}$$

$$\frac{\partial}{\partial \theta_{ij}^{(l)}} J(\theta) = D_{ij}^{(l)} \tag{14}$$

Where,

$\delta^{(L)}$  Represents the error of node in layer  $L$ ,

$\mathbf{a}^{(L)}$  Represents the output of the node in layer  $L$ ,

$\mathbf{y}^{(i)}$  Represents the share price of the  $i^{\text{th}}$  training example,

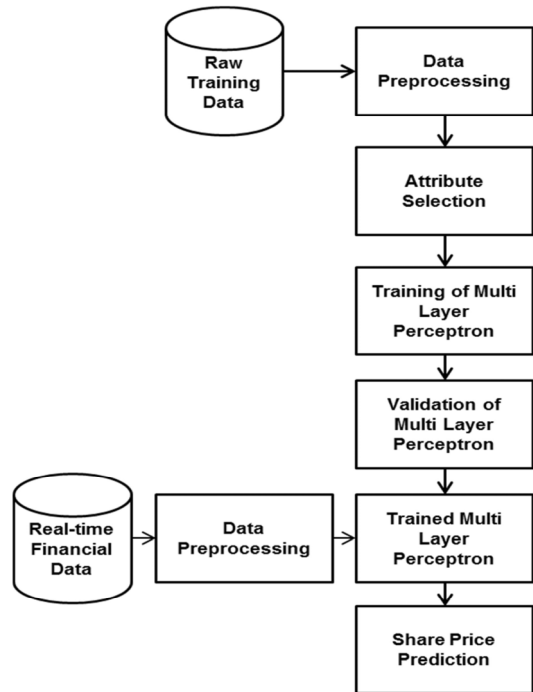
$\Delta_{ij}^{(l)}$  Represent the total error in the  $L^{\text{th}}$  layer,

$D_{ij}^{(l)}$  Represent the partial derivatives of the cost function in the  $L^{\text{th}}$  layer

**ANN Model Training Methodology**

The following <Figure 3> shows the methodology used for developing the artificial neural network for stock price prediction.

Historical financial data about stocks such as financial Ratios, Valuation Ratios and share price is collected from the internet through open source channels and stored in a file format. This raw financial data is preprocessed to eliminate parameters which are not required and clean-up the data for duplicates and missing values. The preprocessed data consists of various financial parameters called as “Attributes” that are assumed to affect the share price, here the machine learning algorithms are used to reduce the feature set and select the most important attributes that affect the share price. The historical data is div-



<Figure 3> Share Price Prediction Methodology

ided in to training and test sets and the ANN is examined using the test data set.

Most machine learning algorithms are designed to learn the most appropriate attributes to use for making their decisions. Because of the negative effect of irrelevant attributes on most machine learning schemes, it is common to precede learning with an attribute selection stage that strives to eliminate all but the most relevant attributes. The best way to select relevant attributes is manually, based on a deep understanding of the learning problem and what the attributes actually mean. However, automatic methods can also be useful. Reducing the dimensionality of the data by deleting unsuitable attributes improves the performance of learning algorithms. It also speeds them up, although this may be outweighed by the computation involved in attribute selection. More

important, dimensionality reduction yields a more compact, more easily interpretable representation of the target concept, focusing the user’s attention on the most relevant variables.

The machine learning algorithm used here is the “CfsSubsetEval” machine learning attribute Subset selector which considers predictive value of each attribute individually, along with the degree of redundancy among them, this algorithm search works by using the “Best First” technique which is based on Greedy hill climbing with backtracking algorithm (Witten et al., 2011a). The attribute selector takes input as the preprocessed financial data and it checks each attribute and its relation to the share price. The attributes that have a direct relation to the share price are selected and these attributes form input to the Share price prediction ANN.

### 3.2. Mathematical Model for Creating Stock Portfolios

The output of the stock price prediction ANN described in the earlier section is the expected stock price  $P_e(S)$ , this expected stock price  $P_e(S)$  is fed

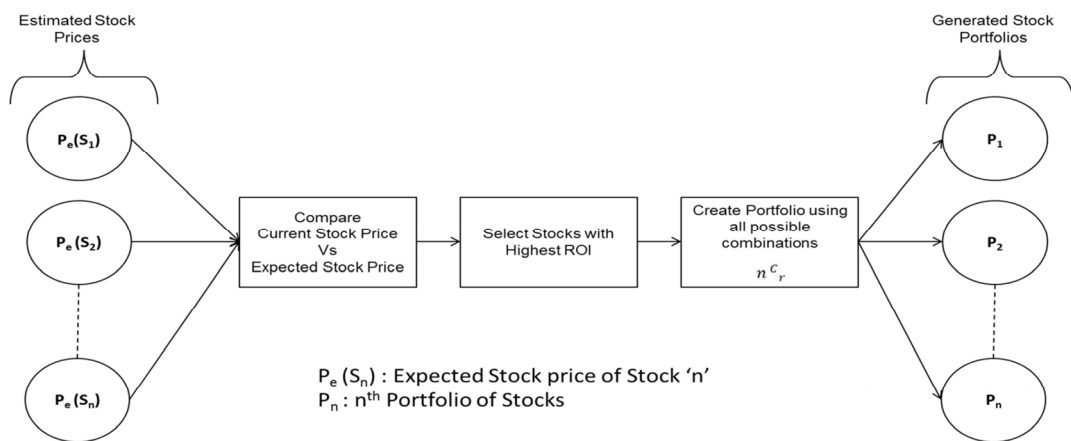
to the mathematical model described below in <Figure 4> to create automatic stock portfolios based on statistical combinations and evaluations.

The mathematical model calculates the difference between the Current Market Price  $P_c(S)$  and expected stock price  $P_e(S)$  and evaluates the net Return on Investment (ROI) for every stock given as an input to the ANN for stock price prediction. The Mathematical Model then selects a subset of stocks with highest ROI. The subset of top performing stocks is used by the model to create portfolios using statistical combinations. These auto generated portfolios are the fed to the portfolio selection algorithms to select the best portfolio based on Risk and Reward analysis.

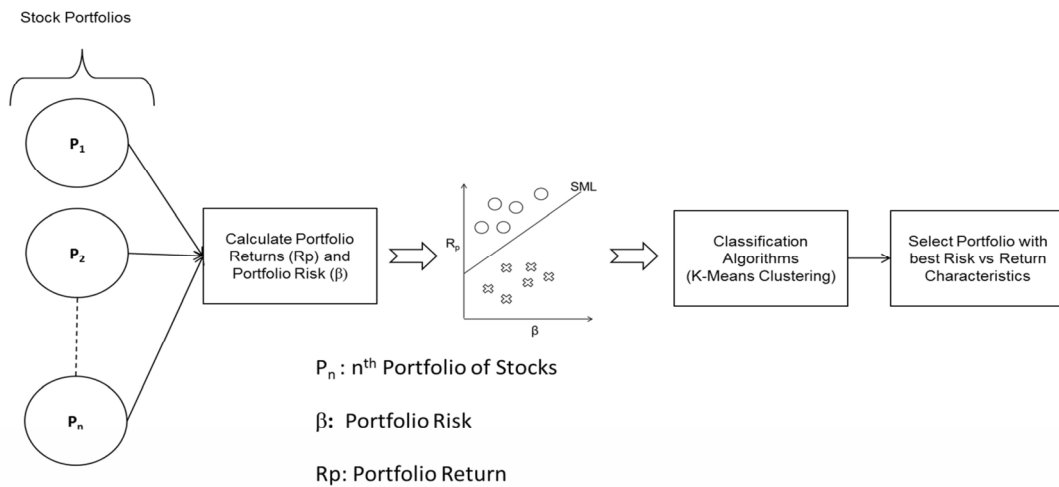
### 3.3. Classification Model for Stock Portfolio Evaluation and Selection

Once the portfolios are created, they need to be evaluated and selected based on the Portfolio Risk ( $R_p$ ) and Portfolio Return ( $\beta$ ).

The Portfolio Return of a combination of ‘n’ stocks is calculated using the below equations:



<Figure 4> Mathematical Model for Creating Stock Portfolios



<Figure 5> Classification Model for Stock Portfolio Evaluation and Selection

$$R_p = \sum_{i=1}^n x_i r_i \quad (15)$$

$R_p$  = Portfolio Return

$x_i$  = Proportion of funds invested in each security

$r_i$  = Expected return of a particular Stock

$n$  = Number of Stocks in the Portfolio

The Systematic Risk of the Portfolio is represented by  $\beta_p$ , The equations for calculating is given below:

$$\beta_p = \sum_{i=1}^n x_i \beta_i \quad (16)$$

$\beta_p$  = Systematic Risk of the Portfolio

$x_i$  = Proportion of funds invested in each security

$\beta_i$  = Systematic Risk of a particular Stock

$n$  = Number of Stocks in the Portfolio

Once the Portfolio Return and Risk are calculated, the data is fed to the Machine learning Logistic Regression Model for classifying the stocks above the Security Market Line (SML), the SML is a line

drawn on a chart that serves as a graphical representation of the capital asset pricing model (CAPM), which shows different levels of systematic, or market, risk of various marketable securities plotted against the expected return of the entire market at a given point in time (Cenesizoglu et al., 2019). Also known as the “characteristic line,” the SML is a visual of the capital asset pricing model (CAPM), where the x-axis of the chart represents risk in terms of beta, and the y-axis of the chart represents expected return. The market risk premium of a given security is determined by where it is plotted on the chart in relation to the SML.

K-means is one of the unsupervised learning algorithms for clustering problem (Kumari et al., 2019). The procedure follows a way to classify a given data set through a certain number of clusters (assume  $k$  clusters). The main idea is to define  $k$  centers, one for each cluster. After we have  $k$  new cluster centroids, this algorithm aims at minimizing cost function know as squared error function given by (Witten et al., 2011b):

$$J(V) = \sum_{i=1}^{c_i} \sum_{j=1}^{c_j} \|x_i - v_j\|^2 \quad (17)$$

Where:

$J(V)$  is the cost function,

$c_i$  is the number of data points in  $i^{\text{th}}$  cluster.

$c_j$  is the number of cluster centers.

$x_i$  is a data point in the cluster data set

$v_j$  is the cluster mean

$\|x_i - v_j\|$  is the Euclidean distance between  $x_i$  and  $v_j$ .

In the last phase of portfolio selection, the stock portfolios above the SML line are selected as they have superior returns for the given Portfolio Risk ( $\beta$ ).

## IV. Data and Results

Based on the ANN Model and methodology, financial data related to the stock prices was collected and used for Attribute selection and training the ANN. The WEKA tool (Bouckaert et al., 2017) was used for all phase of Data Analysis, Machine learning and building the DSS.

### 4.1. Attribute Selection and Analysis

The financial attributes such Investment Valuation Ratios ( $R_v$ ), Profitability Indicator Ratios ( $R_{pf}$ ), Liquidity Measurement Ratios ( $R_l$ ), Debt Ratios ( $R_{db}$ ) and Cash Flow Indicator Ratios ( $R_{cf}$ ) as discussed in section 3.1 of this paper are passed through an attribute selection algorithm “CfsSubsetEval” described in Fig 3.1 of this paper. The selected attributes are as follows:

- EV/EBITDA

- Earnings Yield
- P/E Ratio
- Basic EPS

**EV/EBITDA:** This valuation metric is calculated by dividing a company’s “enterprise value” (EV) by its earnings before interest expense, taxes, depreciation and amortization (EBITDA). This measurement allows investors to assess a company on the same basis as that of an acquirer. As a rough calculation of enterprise value multiple serves as a proxy for how long it would take for an acquisition to earn enough to pay off its costs (assuming no change in EBITDA).

$$EV/EBITDA = \frac{\text{Market Capitalization} - \text{Total Debt} + \text{Cash and Cash Equivalents}}{EBITDA}$$

**Earnings Yield:** The earnings yield refers to the earnings per share for the most recent 12-month period divided by the current market price per share. The earnings yield (which is the inverse of the P/E ratio) shows the percentage of how much a company earned per share. This yield is used by many investment managers to determine optimal asset allocations and is used by investors to determine which assets seem underpriced or overpriced.

$$\text{Earnings Yield} = \frac{\text{Earnings Per Share (EPS)}}{\text{Share Price}}$$

**P/E Ratio:** The price-to-earnings ratio (P/E ratio) is the ratio for valuing a company that measures its current share price relative to its per-share earnings (EPS). The price-to-earnings ratio is also sometimes known as the price multiple or the earnings multiple. P/E ratios are used by investors and analysts to determine the relative value of a company’s shares in

an apples-to-apples comparison. It can also be used to compare a company against its own historical record or to compare aggregate markets against one another or over time. To determine the P/E value, one simply must divide the current stock price by the earnings per share (EPS).

$$\text{P/E Ratio} = \frac{\text{Current Stock Price}}{\text{Earnings Per Share (EPS)}}$$

**Basic EPS:** Earnings per share (EPS) is calculated as a company's profit divided by the outstanding shares of its common stock. The resulting number serves as an indicator of a company's profitability. It is common for a company to report EPS that is adjusted for extraordinary items and potential share dilution. The higher a company's EPS, the more profitable it is considered. The earnings per share metric are one of the most important variables in determining a share's price. By dividing a company's share price by its earnings per share, an investor can see the value of a stock in terms of how much the market is willing to pay for each dollar of earnings. The EPS growth rate is one of the key indicators to evaluate the future profits and trends of a stock.

$$\text{EPS} = \frac{\text{Company Net Earnings}}{\text{Number of common outstanding Shares}}$$

#### 4.2. Data Sources

A subset of large cap companies (50 Companies) listed on the NSE (NSE, 2018) were selected for stock price prediction, the financial data related to stock price was collected from the open source data available on the "Money Control" website ([www.moneycontrol.com](http://www.moneycontrol.com), 2018). The data from the site was collected and stored in a spread sheet for

data cleanup, standardization and later converted to "Comma Separated Values" (.csv) file format to be used by the data analysis tool. Data from last five years on these companies was collected to train the ANN. The data about each stock was collected between the years 2004 to 2018.

#### 4.3. Stock Prediction Results

The Stock Prediction ANN for was trained with the data from Top 50 companies with the largest Market Capitalization from NSE with financial information from years between 2004 and 2018. The below <Table 2> shows the Test evaluation parameters obtained during the testing process using Weka Tool.

The results show that the ANN has a very high level of correlation with the input attributes and predicts the share price with high level of accuracy. The below <Figure 6> shows "Actual" vs "Predicted" stock prices for one of the sample test runs to validate the accuracy of the ANN, which shows high level of correlation between the "Actual" and "Predicted" stock prices.

#### 4.4. Stock Portfolio Generation Results

The stock prediction was done on the Top 50 companies by Market Capitalization listed on the NSE, out of the 50 companies. The top 10 companies with the highest ROI based on the current price and the price predicted for the next quarter were selected for mathematical combinations. Each combination of 5 stocks was done from the list forming each individual portfolio. The mathematical model generated 250 portfolios with 5 companies in each portfolio. The Risk (Beta) and the return were automatically calculated for each portfolio. The portfolio

<Table 1> Sample Snapshot of Financial Data Used for ANN Training and Test

No.	1: EV/EBITDA (X)	2: Earnings Yield	3: P/E Ratio	4: Basic EPS (Rs.)	5: Share Price
	Numeric	Numeric	Numeric	Numeric	Numeric
1	16.1	0.05	21.72	131.15	2847.92
2	15.09	0.05	20.26	120.04	2431.75
3	15.94	0.05	21.47	117.11	2514.75
4	18.93	0.04	25.9	98.31	2545.93
5	16.41	0.04	22.6	94.15	2127.55
6	18.37	0.04	24.11	65.22	1572.35
7	16.03	0.05	20.87	55.95	1167.5
8	24.68	0.03	30.62	38.61	1182.42
9	21.84	0.04	27.21	28.71	781.11
10	9.23	0.09	11.25	47.99	540.06
11	14.44	0.06	17.61	46.07	811.21
12	26.56	0.03	32.07	38.39	1231.13
13	28.03	0.03	34.48	55.53	1914.49
14	28.05	0.03	37.55	38.15	1432.51
15	10.9	0.06	16.66	53.08	884.05
16	10.17	0.07	26.55	49.77	1321.38
17	8.89	0.08	12.35	84.66	1045.2
18	8.54	0.09	11.7	70.25	821.62
19	8.77	0.07	13.62	68.05	926.84
20	6.57	0.08	12.04	64.82	780.4
21	6.63	0.08	12.44	61.21	761.67
22	9.21	0.06	17.55	62.0	1088.05
23	12.12	0.05	23.15	49.64	1149.24
24	11.31	0.06	17.24	97.29	1677.6
25	14.93	0.06	17.05	134.21	2288.62
26	10.55	0.06	16.64	85.73	1426.56
27	8.72	0.08	13.51	65.1	879.43
28	6.4	0.1	10.87	53.88	585.86
29	8.72	0.07	16.13	36.27	584.92
30	17.53	0.04	27.78	9.99	956.00

<Table 2> Test Evaluation Parameters for Stock Price Prediction

Test Evaluation Parameter	Value
Correlation coefficient	0.9968
Mean absolute error	86.8413
Root mean squared error	127.3714
Relative absolute error	7.77%
Root relative squared error	8.77%

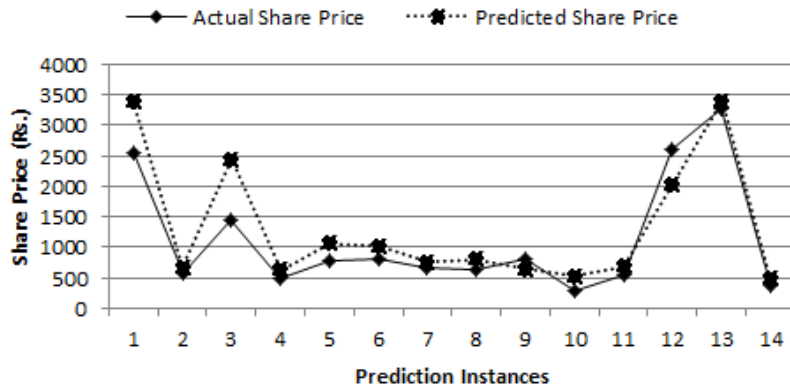
selection was done using K-Means clustering machine learning algorithm with highest Return and lowest risk.

Based on the above <Figure 7>, the portfolios

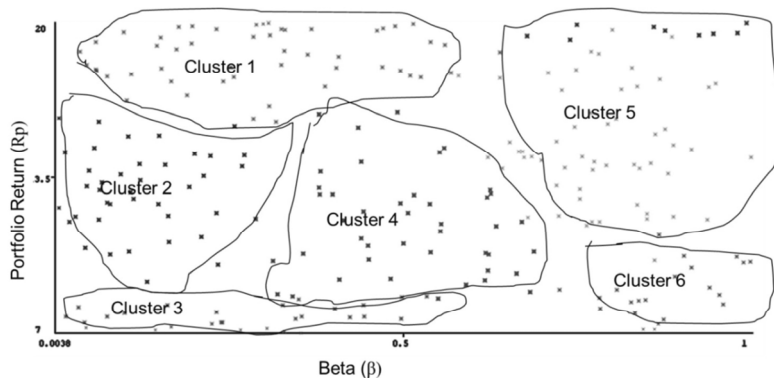
in Cluster 1 are selected since they have the highest return and the lowest risk profile.



## Actual vs Predicted Share Price



<Figure 6> Test Run of "Actual" and "Predicted" Stock Prices



<Figure 7> Portfolio Generation and Clustering by K-Means Algorithm

## V. Findings and Implications

### 5.1. Major Findings of the Study

The major findings of this study which consisted of select stocks listed on the NSE are as follows:

- It is possible to train the ANN models with past financial data and predict future prices with a hybrid model of intrinsic and exogenous feature set
- Choosing the intrinsic feature set for ANN using machine learning based attribute selector increases the prediction accuracy
- It is possible to generate a portfolio of top performing stocks based on machine learning un-supervised Classification algorithms.
- The classification algorithms are very effective in performing the risk analysis and selecting stocks with lowest risk and highest returns.
- The financial decision support systems are a valuable source of tools for individual investors

as they enable them to cut losses and reduce risk in their trading operations

## 5.2. Implications of the Study

The implications of this study for predicting the Stock Prices, Creation of Automatic Stock Portfolios and assisting individual investors with decision support are multi-fold:

- The Intrinsic features ensure that predicted stock prices are more closer to the real-time market value eventually may contribute to a more efficient financial system.
- Automatic Generation of stock portfolios will become efficient as the selection will be based on machine learning risk analysis and reduces the human error
- Financial Decision Support Systems can revolutionize virtually every aspect of financial and investment decision making. Financial firms worldwide can employ neural networks to tackle difficult tasks involving intuitive judgement or requiring the detection of data patterns which elude conventional analytic techniques.
- The Financial Decision Support Systems benefits individual investors with little knowledge of the dynamics involved in stock markets, any reasonable prediction can enable common investors to invest in stock markets and enable the growth of the economy.

## 5.3. Limitations of the Study

The model developed in this paper for stock price prediction and portfolio selection is based on funda-

mental attributes which will represent the intrinsic value of the stock and does not represent the real-world market value of the stock. The market value will eventually converge to the intrinsic value of the stock on a long term basis. Below are the limitations of this study:

- The stock predictions of this model cannot be used for Day-to-Day prediction and not useful for day traders
- The stock price predictions and Portfolio selections are only valid for medium term and long term investments which are more than a quarter or more

## 5.4. Conclusions

The authors have developed a Machine learning Model using ANNs for stock price prediction using fundamental attributes of a stock in contrast to the existing techniques that use technical attributes (Khaidem et al., 2016) of a stock. The authors also have investigated the traditional models of stock portfolio selection which are based on statistical and quantitative analysis (Dipietro, 2019) (Text, 2018), the authors have recognized this gap in the literature and have developed a machine learning clustering model to build a portfolio of top performing stocks with premium returns and the least risk compared to other portfolio clusters. The stock prediction ANN and clustering model for portfolio selection have been subjected to real world data to validate the models and the results have provided to be accurate enough to deploy these models in the financial domain for accurate decision support.

## &lt;References&gt;

- [1] Agrawal, S., Thakkar, D., Soni, D., Bhimani, K., and Patel, C. (2019). Stock market prediction using machine learning techniques. *Data Collection Feature Extraction Data Normalization Training Output*, 5(2), 1099-1103.
- [2] Anbalagan, T., and Maheswari, S. U. (2014). Classification and prediction of stock market index based on Fuzzy Metagraph. *Procedia Computer Science*, 47(C), 214-221. <https://doi.org/10.1016/j.procs.2015.03.200>
- [3] Arik, S., Eryilmaz, S. B., and Goldberg, A. (2014). *Supervised classification-based stock prediction and portfolio optimization*. Retrieved from <http://arxiv.org/abs/1406.0824>
- [4] Banik, S., Khodadad Khan, A. F. M., and Anwer, M. (2014). Hybrid machine learning technique for forecasting dhaka stock market timing decisions. *Computational Intelligence and Neuroscience*, 2014. <https://doi.org/10.1155/2014/318524>
- [5] Bouckaert, R. R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A., and Scuse, D. (2017). *WEKA Manual for Version 3-8-2*.
- [6] Cenesizoglu, T., Papageorgiou, N., Reeves, J. J., and Wu, H. (2019). An analysis on the predictability of CAPM beta for momentum returns. *Journal of Forecasting*, 38(2), 136-153. <https://doi.org/10.1002/for.2552>
- [7] Cheng, C., and Chen, Y. (2007). Fundamental analysis of stock trading systems using classification techniques. *Proceedings of the Sixth International Conference on Machine Learning and Cybernetics*, Hong Kong, 19-22 August 2007, 19-22.
- [8] Chong, E., Han, C., and Park, F. (2017). Deep learning networks for stock market analysis and prediction. *Expert Systems with Applications*, 83(April), 187-205. <https://doi.org/10.1016/j.eswa.2017.04.030>
- [9] Dase, R. K., and Pawar, D. D. (2010). Application of artificial neural network for stock market predictions: A review of literature. *International Journal of Machine Intelligence*, 2(2), 14-17. <https://doi.org/10.9735/0975-2927.2.2.14-17>
- [10] Di Persio, L., and Honchar, O. (2016). Artificial neural networks architectures for stock price prediction: Comparisons and applications. *International Journal of Circuits, Systems and Signal Processing*, 10, 403-413. <https://doi.org/10.1676/09-204.1>
- [11] Diakoulakis, I. E., Koulouriotis, D. E., and Emiris, D. M. (2018). A Review of Stock Market Prediction Using Computational Methods. In: Kontoghiorghes E.J., Rustem B., Siokos S. (eds) *Computational Methods in Decision-Making, Economics and Finance*. Applied Optimization, vol 74. Springer, Boston, MA. [https://doi.org/10.1007/978-1-4757-3613-7\\_20](https://doi.org/10.1007/978-1-4757-3613-7_20)
- [12] Dipietro, D. M. (2019). Alpha cloning: Using quantitative techniques and SEC 13f data for equity portfolio optimization and generation. *The Journal of Financial Data Science Fall 2019*, 1 (4), 159-171. <https://doi.org/10.3905/jfds.2019.1.008>
- [13] Dunne, M. (2017). Stock market prediction declaration of originality. *Dept of Computer Science, University College Cork*, 1(1), 10.
- [14] Flechas Chaparro, X. A., de Vasconcelos Gomes, L. A., and Tromboni de Souza Nascimento, P. (2019). The evolution of project portfolio selection methods: from incremental to radical innovation. *Revista de Gestão*, 26(3), 212-236. <https://doi.org/10.1108/rega-10-2018-0096>
- [15] Huang, Y., Fernando Capretz, L., and Ho, D. (2019). Neural network models for stock selection based on fundamental analysis. *IEEE*, Canada, (May), 1-4.
- [16] Ican, Ö., and Çelik, T. B. (2017). Stock market prediction performance of neural networks: A literature review. *International Journal of Economics and Finance*, 9(11), 100. <https://doi.org/10.5539/ijef.v9n11p100>
- [17] Khaidem, L., Saha, S., and Dey, S. R. (2016). *Predicting the direction of stock market prices using random forest*, 1-20. Retrieved from <http://arxiv.org/abs/1605.00003>
- [18] Kumari, S. K., Kumar, P., Priya, J., Surya, S., and

- Bhurjee, A. K. (2019). Mean-value at risk portfolio selection problem using clustering technique: A case study. *AIP Conference Proceedings*, 2112(June). <https://doi.org/10.1063/1.5112363>
- [19] Kusuma, R. M. I., Ho, T.-T., Kao, W.-C., Ou, Y.-Y., and Hua, K.-L. (2019). *Using deep learning neural networks and candlestick chart representation to predict stock market*, 1-13. Retrieved from <http://arxiv.org/abs/1903.12258>
- [20] Leong, Y. C., and Zaki, J. (2018). Unrealistic optimism in advice taking: A computational account. *Journal of Experimental Psychology: General*, 147(2), 170-189. <https://doi.org/10.1037/xge0000382>
- [21] Malagrino, L. S., Roman, N. T., and Monteiro, A. M. (2018). Forecasting stock market index daily direction: A Bayesian Network approach. *Expert Systems with Applications*, 105, 11-22. <https://doi.org/10.1016/j.eswa.2018.03.039>
- [22] Malkiel, B. G. (2003). Efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59-82.
- [23] Montmarquette, C., and Viennot-briot, N. (2012). *Econometric models on the value of advice of a financial advisor*. Centre Interuniversitaire de Recherche En Analyse Des Organisations.
- [24] NSE. (2018). NSE. Retrieved from <https://www.nseindia.com/products/content/equities/equities/equities.htm>
- [25] Pang, X., Zhou, Y., Wang, P., Lin, W., and Chang, V. (2018). An innovative neural network approach for stock market prediction. *Journal of Supercomputing*, (January), 1-21. <https://doi.org/10.1007/s11227-017-2228-y>
- [26] Putra, A. I. L. W., Putra, A. D., Dewi, M. S., and Radianto, D. O. (2019). Differences in intrinsic value with stock market prices using the price earning ratio (per) approach as an investment decision making indicator (Case study of manufacturing companies in Indonesia period 2016-2017). *Aptisi Transactions on Technopreneurship*, 1(1), 82-92. <https://doi.org/10.34306/att.v1i1.61>
- [27] Pyo, S., Lee, J., Cha, M., and Jang, H. (2017). Predictability of machine learning techniques to forecast the trends of market index prices: Hypothesis testing for the Korean stock markets. *PLoS ONE*, 12(11), 1-17. <https://doi.org/10.1371/journal.pone.0188107>
- [28] Qiu, M., and Song, Y. (2016). Predicting the direction of stock market index movement using an optimized artificial neural network model. *PLoS ONE*, 11(5), 1-11. <https://doi.org/10.1371/journal.pone.0155133>
- [29] Report, T. (2016). Automated stock market trading system using machine learning. *Automated Stock Market Trading System*, (May 2015). <https://doi.org/0.13140/RG.2.1.1998.3520>
- [30] Rudin, C. (2012). *A profitable approach to security analysis using machine learning: an application to the prediction of market behavior following earnings Reports*. 15.097 Prediction: Machine Learning and Statistics (MIT-OCW), 1-22. Retrieved from [https://ocw.mit.edu/courses/sloan-school-of-management/15-097-prediction-machine-learning-and-statistics-spring-2012/projects/MIT15\\_097S12\\_proj2.pdf](https://ocw.mit.edu/courses/sloan-school-of-management/15-097-prediction-machine-learning-and-statistics-spring-2012/projects/MIT15_097S12_proj2.pdf)
- [31] Selvamuthu, D., Kumar, V., and Mishra, A. (2019). Indian stock market prediction using artificial neural networks on tick data. *Financial Innovation*, 5(1). <https://doi.org/10.1186/s40854-019-0131-7>
- [32] Song, I. (2014). New quantitative approaches to asset selection and portfolio construction. ProQuest Dissertations and Theses, 213. <https://doi.org/10.13005/ojc/290419>
- [33] Sung, M.-C., Ma, T., Hsu, M.-W., Johnson, J. E. V., and Lessmann, S. (2016). Bridging the divide in financial market forecasting: Machine learners vs. financial economists. *Expert Systems with Applications*, 61, 215-234. <https://doi.org/10.1016/j.eswa.2016.05.033>
- [34] Text, F. (2018). *Accern and truerisk labs announce the next generation of unique machine-learning trading signals and portfolios*, (July), 1-3.
- [35] Tsai, C.-F. and Wang, S.-P. (2009). Stock price forecasting by hybrid machine learning techniques. *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 1. Retrieved from

- [http://www.iaeng.org/publication/IMECS2009/IMECS2009\\_pp755-760.pdf](http://www.iaeng.org/publication/IMECS2009/IMECS2009_pp755-760.pdf)
- [36] Wen, M. I. N., Li, P., Zhang, L., and Chen, Y. A. N. (2019). Stock market trend prediction using high-order information of time series. *IEEE Access*, 7, 28299-28308. <https://doi.org/10.1109/ACCESS.2019.2901842>
- [37] Witten, I. H., Frank, E., and Hall, M. A. (2011a). *Data Mining Practical Machine Learning Tools and Techniques* (3rd ed.). Elsevier Inc.
- [38] Witten, I. H., Frank, E., and Hall, M. A. (2011b). *Data Mining Practical Machine Learning Tools and Techniques*, 3rd Editio, 174-176.
- [39] [www.moneycontrol.com](http://www.moneycontrol.com). (2018). MoneyControl.com. Retrieved from <http://www.moneycontrol.com>
- [40] Yong, C. C., and Taib, S. M. (2009). Designing a decision support system model for stock investment strategy. *Proceedings of the World Congress on Engineering and Computer Science 2009 Vol I, WCECS 2009*, October 20-22, 2009, San Francisco, USA.
- [41] Zhang, L., Aggarwal, C., and Qi, G.-J. (2017). Stock price prediction via discovering multi-frequency trading patterns. *KDD '17: Proceedings of the 23rd ACM*, 2141-2149. <https://doi.org/10.1145/3097983.3098117>

◆ About the Authors ◆

---



**Sandeep Patalay**

Sandeep Patalay is a Ph.D research scholar at Vignan's University and pursuing research in the area of Financial Decision Support Systems. He has a master's degree in Finance from Acharya Nagarjuna University, Guntur, India and Bachelor's degree in Electronics Engineering from Jawaharlal Nehru Technological University, Hyderabad, India. He works at Tata Consultancy Services, Hyderabad, India and leads the Centre of Excellence for IoT technologies. He has over 16 Years of experience in the area of Analytics and Decision Support Systems.



**Madhusudan Rao Bandlamudi**

Dr. Madhusudan Rao Bandlamudi works at Vignan's University as a Profosser of Management studies. He has over 30 years of teaching experience and 10 years of Industry experience. He is currently guiding Ph.D research students in the areas of Financial Management, Operations Management and Entrepreneurship

---

Submitted: April 1, 2019; 1st Revision: September 7, 2019; Accepted: October 4, 2019