Capital Asset Pricing Model and Artificial Neural Networks: A Case of Pakistan’s Equity Market

Usman Ayub
Department of Management Sciences,
COMSATS University Islamabad, Pakistan
Email: usman_ayub@comsats.edu.pk

Muhammad Naveed Jan
Department of Management Sciences,
COMSATS University Islamabad, Pakistan
Email: naveedjan@ciit.net.pk

Muhammad Asim Afridi
Department of Management Sciences,
COMSATS University Islamabad, Pakistan
Email: asim.afridi@gmail.com

Imran Abbas Jadoon
Department of Management Sciences,
COMSATS University Islamabad, Pakistan
Email: imran.jadoon@comsats.edu.pk

Abstract:
Artificial Neural Networks (ANN) approach is a relatively new and promising field of the prediction of stock price behavior. Neural networks approach is a mathematical model, flexible enough to accommodate both linear and non-linear aspect of stock returns. This paper applies the ANN to asset pricing models. It is found that the optimum number of neurons does not follow some mathematical rule rather it is based on the presentiment of the researcher to apply an exhaustive search for the number of optimum neurons. Another important finding is that the difference of errors between the testing and training dataset is minimum and the networks are not suffering from the over-fitting phenomenon. The predicted value of high beta portfolios is better than the low beta and mid beta portfolios. This finding reinforces the investment principle that the market compensates the high-risk portfolios more than other classes. The paper concludes that the proposed model achieves a significant improvement in the return on investment and the investors can magnify their profitability.

Keywords: Artificial Neural Networks; Asset Pricing Models; Stock Markets

I. Introduction
Forecasting the stock returns is an exciting and arguable issue among the finance researchers and the asset pricing theorists. Although Cabrera et al. (2011) consider the subject of forecasting merely a data snooping exercise, the vast repository of prediction studies provide a solid background of the successful estimation of stock markets returns. Studies have provided meaningful evidence that predicting the stock market returns is now an attainable goal(Ang & Bekaert, 2006a; Goyal & Welch, 2003; Rapach&Wohar,
2006). These studies document that the prospects of huge economic benefits and wealth maximization of the investors are now all-time high with the application of innovative financial modeling. The financial modeling of the stock returns possesses peculiar characteristics (Keim & Stambaugh, 1986; Fama & French, 1988; Harvey, 1995; Schiller & Campbell, 1998; Campbell & Shiller, 2005). Firstly, the stock prices and returns are predictable by a set of financial, technical and economic variables. Secondly, a linear relationship exists between the risk and returns. Thirdly, the linear forecasting models are a better and convenient fit for the stock markets. On the basis of these assumptions, the use of linear predictive models, in most of the empirical finance research is simple, but the related econometric problems of forecasting accuracy are numerous. Ang & Bekaert(2006b) have pointed out that these models suffer from inherent limitation to capture the impact of the nonlinear nature of the stock market data.

Similarly, Campbell & Shiller(2005) have documented that the actual relationship between the predictor variables and long-term stock returns is nonlinear, and the use of linear regression might produce bearish results. Another problem with the linear predictive models, in the long run, is the use of overlapping data which causes the error terms to be strongly correlated. Studies have shown that linear prediction models perform poorly in sample and out sample datasets in nonlinear markets and therefore are of little use for practical purposes (Butler et al., 2005; Campbell & Thompson, 2007; Goyal & Welch, 2003). McMillan(2005) has found that the global financial markets behave in a nonlinear manner and therefore the application of nonlinear technique will capture this turbulent phenomenon accurately. The artificial neural networks is a nonlinear and non-parametric technique that can accurately capture the nonlinear nature of stock market data. The mechanism and algorithm of artificial neural networks closely resemble the human brains in decision-making(Guotai et al. 2017). For example, when the market forces cause a change of 5% or less in the required rate of returns of the investors, they may ignore it. If this trend continues and the rates of returns change above the threshold level of the investors, they positively frame their reaction and alter their decisions accordingly. The ANN has an established ability to define these situations in financial markets and other conditions (Kamruzzaman & Sarker, 2003). Empirical finance is using this epistemology of artificial neural networks in decision making.

The forecasting capability of ANN is evidenced by a vast repository of literature. Some recent studies provide evidence of the forecasting ability of ANN in stock markets(Rekik, Hachicha, & Boujelbene, 2014; Qiu, Song, & Akagi, 2016; Tkáč & Verner, 2016; Guan, Dai, Zhao, & He, 2018). These studies have some major limitations. Firstly, the selection of input variables based on the researcher’s choice which are not supported by asset pricing theory. Secondly, the interpretation of the index movements in terms of investors returns is possible in those markets where index derivatives are available. Thirdly, there is limited options of datasets usage. Fourthly, these studies are useful for the individual stocks and day traders only. Finally, most of these studies utilize the generic software of ANN with limited options. The long-term portfolio investors, on the other hand, mainly rely on the composite factors of asset pricing models for forecasting their returns. This difference in the investment strategies of the day traders and long-term investors is the main drawback of ANN literature on forecasting in stock markets. The composite factor of market return of the pioneering asset pricing model CAPM (Capital Asset Pricing model) has an inbuilt explanation for major variations in stock returns. This model is based on the pioneering work of Markowitz(1952) and draws
its underlying assumption from the modern portfolio theory with additional assumption of unlimited borrowing, lending and permission for short sales (Sharpe, 1964). The model is a cornerstone of pricing risky assets in financial economics and suggests that the excess stock returns are related to the systematic risk of the market. According to Fama & French(2004), the CAPM is not only strong theoretically, but the mathematical derivation is also very appealing when the expected rate of return and risk are calculated.

The predictive performance of the CAPM in conditional and unconditional form has been investigated by many studies. These studies use linear techniques to assess the predictive performance of CAPM and its variants. The results of these studies are mixed regarding the predictive power of traditional and conditional asset pricing models. The objective of this study is to investigate the forecasting ability of ANN from broad horizons including the distribution of datasets, the number of neurons, the application of various training functions and their parameters by using capital asset pricing model under rolling window schemes. This study is expected to solve the major concern of the investors to overcome the uncertainty of the equity markets in emerging countries and generate significant economic value. The remaining paper is organized as follows. Section 2 provides literature review. Section 3 presents data and theoretical framework. Section 4 explains empirical results. Final section concludes the paper.

II. Literature Review

The rise and fall of the market prices of the risky assets are interpreted by investors in different ways. They demonstrate their reaction in the form of buying and selling decisions of the securities. This response points towards the learning and generalization capability of human beings(Guan et al., 2018). Financial modeling is based on this psychology of the investors and the key players in financial markets, although, have full access to all types of information, they still have to learn the appropriate reaction on the spot. Artificial neural networks are believed to possess the ability to mimic human psychology (Guotai et al., 2017) and can be employed as a replacement for rational thinking in the financial markets. The salient characteristics of this technique include high resilience to deal with uncertainty, imprecision, and management of highly noisy datasets(Rekik et al., 2014). This technique is taking over the traditional linear and nonlinear techniques and can be employed to predict stock prices, portfolio returns, directional changes in the stock market index. Carvalhal& Ribeiro(2008) compare the predictive power of ANN with three traditional forecasting models, i.e. Random walk, Autoregressive moving averages, and generalized autoregressive conditional heteroskedasticity models. The study provides substantial evidence that the forecasting performance of ANN surpasses other techniques in predicting the indices.

Mostafa (2010) studies the forecasting ability of various architectures of ANN and finds that ANN models are valuable tools in predicting stock exchange movements in developing markets. ANNs also perform well to estimate the directional movement of the index. Kara et al. (2011) suggest that the neural network forecasts the direction of the index with 75.74% accuracy, while the other model demonstrates 71.52% accuracy. Similarly, Kumar et al.(2011) have shown that ANNs resist the noisy environments like stock market data and demonstrate a high tolerance to fuzziness. These characteristics make ANNs suitable for the use in the area of stock market modeling. The study concludes that ANN is ideal for making informed investments decisions in the share market.
ANNs also return promising forecasting results in the long run. Al-Jarrah et al. (2011) report that the long-term predictive accuracy of Artificial neural works is 80%. This study suffers from many weaknesses including the use of raw data in artificial neural networks. This drawback has probably reduced the forecasting accuracy of this research. Idowu et al. (2012) point out that although ANNs do not allow perfect estimations on volatile data such as the stock exchange market, they indeed provide closer results to the real ones compared with other techniques. Similarly, Maknickienë & Maknickas (2013) suggest a forecasting model for foreign exchange markets based on artificial neural works. The findings of the study consider the ANN as a useful tool for investors, making his portfolio construction process more profitable.

The application of ANN in Pakistan’s equity is taken up by many scholars. The study of Fatima & Hussain (2008) is perhaps the first contribution in this regard. This paper compares the forecasting performance of ANN, ARIMA, and GARCH and assesses the prediction performance by using forecast mean squared error. The study affirms ANN to be a better model for predicting the stock returns in this volatile market. Haider & Nishat (2009) apply the feedforward neural networks to the stock values of a single firm. Particle swarm optimization (PSO) is applied for portfolio construction. The study successfully achieves excess returns as compared to the market. Iqbal (2013) undertakes the application of several neural network models to predict the returns of a single stock in Pakistan stock exchange and suggests that the neural networks have successfully predicted the returns with more accuracy. All these studies have relied on the use of technical analysis but have ignored the established finance theory.

The first attempt to evaluate the forecasting ability of the asset pricing model, using artificial neural networks is taken up by Cao et al. (2011). The study compares the forecasting performance of single factor dynamic capital asset pricing model, Fama & French three factors model, and ANN. The study suggests that ANN predicts the stock prices more accurately on the asset pricing models. Gokgoz & Sezgin-Alp (2014) evaluates the predictive performance of the asset pricing models using artificial neural networks. The findings suggest that ANN has successfully predicted the future returns of all the indices with a minimum error both on in sample and out sample dataset. Similarly, Aldaarmi et al. (2015) investigate the predictive ability of Fama and French three factors model and value-based management model (VBM) using artificial neural networks. The results suggest that the difference between the actual and predicted portfolio return is significant.

The present study is an endeavor to benefit from the massive computational power of the current times and applies a pure engineering and computer science based tools and pave the way for interdisciplinary research. The existing studies on Pakistan's equity market including Iqbal et al. (2013), Haider and Nishat (2010), and Fatima and Hussain, (2008) have looked at the limited aspects of this subject. The present study has very broad testing horizons regarding the network testing parameters, time span, input out selection, investment theory, and designing of various programs. The application of well-known investment methods in forecasting the stock prices and paving the way for robust investment strategies in an emerging market is a significant contribution to the existing body of knowledge.
III. Data and Methodology

The CAPM states that asset returns are described by its systematic risk relative to the market return. This model incorporates the market return factor of $R_m$ which needs to construct for the proposed model of this study. The CAPM has the following specification:

$$R_p = R_f + \beta_p [R_m - R_f] \quad \ldots \ldots (1)$$

where

$$\beta_p = \frac{\sigma_{pm}}{\sigma_m^2} = \frac{\text{cov}(r_p, r_m)}{\text{var}(r_m)} \quad \ldots \ldots (2)$$

Likewise, a proxy for market portfolio used in CAPM based models is taken as the closing PSX index as the standard procedure following Ayub et al. (2015). Stock prices and PSX index are converted into returns using log difference returns. The impact of dividends is emulated by the stock prices (Ahmed & Javid, 2008). We divide the sample Firms into 30 portfolios based on high, mid and low betas of the sample Firms. In order to forecast in CAPM and ANN frameworks, this study uses ANN with Back Propagation training method as recommended by Sharma (2013). Like any other forecasting tool, the artificial neural networks require the declaration of parameters before putting into operations using appropriate hidden layer/s, neurons and their weights. The study uses supervised learning like training under an instructor, the output is shown to the network, and the prediction is updated according to this actual value (Kamruzzaman & Sarker, 2003). The sigmoid squashing function is more appropriate for a neural network system. Firstly, due to its nonlinear nature, it can successfully capture the underlying relationship between the dependent and independent variables. Secondly and most importantly its differentiation power reduces the error to a minimum level. (Kao et al., 2013). Furthermore, we select the Levenberg-Marquardt (LM) method of training the network as most of the stock market research utilizes the backpropagation architecture along the LM method (Burney et al., 2005). CAPM is interpreted in terms of ANN is described by the following expression

$$R_p = G \left( a + \sum_{j=1}^{h} a_j \right) + F \left( \beta_{0j} + \beta_{1j} (R_{PSX} - R_f) \right) \quad \ldots \ldots (3)$$

The above equation is the translation of equation (1) into ANN framework. We divide the data into a series of combinations to achieve the optimal datasets. The first set is divided into 60-20-20 combination for training, validation and training set. After this first data set distribution, a five percent variation is allowed to the dataset. The generated datasets are presented in tables. These tables show that the dataset contains the combinations of 60-20-20, 65-15-20, 65-10-15, 70-10-20, 70-15-15, 70-20-10, 75-10-15, 75-15-10, 80-10-10, 85-05-10, 85-10-05, and 90-05-05. The detailed network paradigm for this study is executed by initialization architecture and includes three layers, i.e., an input layer, a hidden layer, and an output layer. The input layer utilizes one neuron and the logistic function to assign the appropriate weights to the input variables. The single neurons represent the market returns factor of CAPM. The logistic function assigns the optimum weights to these variables and the output from the input layer is used by the hidden layer. The maximum limit of neurons in the hidden layer is placed at 50, and the
output layer consists of a single neuron in the form of the desired target output. These neurons work in a series, and the programs compile the best results for each neuron under each combination of the dataset. The default performance function for feedforward networks is mean square error—the average squared error between the network outputs and the target outputs $R_t$. The minimum MSE score is chosen as the best lag point for each network system. The mathematical equation of MSE is given below.

$$MSE = \frac{1}{N} \sum_{t=1}^{N} (R_t - \hat{R}_t)^2 \quad \text{(4)}$$

Where $R_t$ and $\hat{R}_t$ are the actual returns and forecasted returns, respectively, and $N$ is the size of the testing dataset. The relationship between the dependent variable (actual or monthly target returns) and the returns (outputs) produced by the neural network system is described as the R-value (regression value). If the R-value is substantial, then MSE value is much smaller than the mean target variance. This indicates that the neural network has successfully managed to model most of the variations in the input-output transformation. An R-value of 1 or 100% shows a close dependence while an R-value of 0 means poor or random relationship. The second set of instructions in Matlab calculates the minimum error between the actual and predicted portfolio returns by 16 datasets and 1-50 neurons. This program is based on a simple step ahead observation reading scheme and generates 800 results for a single portfolio. The third program processes the data generated by the second code in the form of a graph to arrive at the optimal number of neurons and the best dataset. Finally, the rolling windows scheme is implemented to calculate the actual and predicted values by ANN.

IV. Results and Discussion

The study uses data for manufacturing sector for the period January 2000 to December 2015. The data is taken from Pakistan Stock Exchange and Data Stream. Other sources are the Companies financial reports and the State Bank of Pakistan. The major goal of this study is to look for the best ANN architecture of CAPM, which can predict the stock returns very close to the actual returns. The experimentation by the Matlab program generated a total of 800 MSE results for a single portfolio based on 16 data combinations and 50 neurons. Each model utilized 180 monthly returns for the 30 portfolios with one time step. In this way, the network predicted fifty values for one combination of the dataset, and the difference between the anticipated returns and realized returns is estimated in the form of mean squared error. Initially, the models display only the best results for all the three performance measures. The experimentation has been designed in such a way that the system displays the best MSE results as an average of training, validation, and testing.

The experimentation produced a total of 24000 models for CAPM. Table 1 represents the results of the Neural Network models of all the datasets based on CAPM factor. Table 1 shows that the MSE results for the datasets 60-20-20, 65-15-20, 65-20-15, 70-15-15, 70-10-20, and 70-20-10, are almost identical: ranging from 0.0063 to 0.773 for the whole range of neurons. These are the average MSE scores for all the thirty portfolios. The MSE score for the dataset 75-10-15, 75-15-10, 80-15-05 and 75-20-05, are identical and falling in between 0.0064 to 0.77 for all the neurons of the models. The errors reported by other datasets, i.e., 80-10-10, 85-05-10, 85-10-05 and 90-05-05 are ranging from 0.0066 to 0.774.
Table 1: MSE Statistics of all Datasets under 1-50 Neurons (N)

<table>
<thead>
<tr>
<th>CAPM</th>
<th>N1</th>
<th>N5</th>
<th>N10</th>
<th>N15</th>
<th>N20</th>
<th>N25</th>
<th>N30</th>
<th>N35</th>
<th>N40</th>
<th>N45</th>
<th>N50</th>
</tr>
</thead>
<tbody>
<tr>
<td>60-20-20</td>
<td>0.0063</td>
<td>0.0065</td>
<td>0.0082</td>
<td>0.0276</td>
<td>0.1039</td>
<td>0.2207</td>
<td>0.258</td>
<td>0.3598</td>
<td>0.7166</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>65-15-20</td>
<td>0.0064</td>
<td>0.0065</td>
<td>0.0085</td>
<td>0.0397</td>
<td>0.2131</td>
<td>0.2658</td>
<td>0.3767</td>
<td>0.7207</td>
<td>0.7304</td>
<td>0.7869</td>
<td></td>
</tr>
<tr>
<td>60-20-15</td>
<td>0.0063</td>
<td>0.0065</td>
<td>0.0082</td>
<td>0.0276</td>
<td>0.1039</td>
<td>0.2207</td>
<td>0.258</td>
<td>0.3598</td>
<td>0.7166</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>70-10-20</td>
<td>0.0065</td>
<td>0.0067</td>
<td>0.0084</td>
<td>0.0282</td>
<td>0.1039</td>
<td>0.2208</td>
<td>0.2895</td>
<td>0.3604</td>
<td>0.7177</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>70-15-15</td>
<td>0.0064</td>
<td>0.0065</td>
<td>0.0085</td>
<td>0.0281</td>
<td>0.1039</td>
<td>0.2208</td>
<td>0.2894</td>
<td>0.3597</td>
<td>0.7177</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>70-20-10</td>
<td>0.0064</td>
<td>0.0065</td>
<td>0.0083</td>
<td>0.0276</td>
<td>0.1039</td>
<td>0.2207</td>
<td>0.258</td>
<td>0.3597</td>
<td>0.7177</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>75-15-10</td>
<td>0.0066</td>
<td>0.0067</td>
<td>0.0085</td>
<td>0.0282</td>
<td>0.104</td>
<td>0.2208</td>
<td>0.2897</td>
<td>0.3604</td>
<td>0.7177</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>75-20-5</td>
<td>0.0064</td>
<td>0.0065</td>
<td>0.0083</td>
<td>0.0276</td>
<td>0.1039</td>
<td>0.2548</td>
<td>0.2548</td>
<td>0.3597</td>
<td>0.7177</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>80-15-10</td>
<td>0.0067</td>
<td>0.007</td>
<td>0.0088</td>
<td>0.0295</td>
<td>0.1044</td>
<td>0.2212</td>
<td>0.29</td>
<td>0.3617</td>
<td>0.719</td>
<td>0.7262</td>
<td>0.7771</td>
</tr>
<tr>
<td>80-10-10</td>
<td>0.0066</td>
<td>0.0067</td>
<td>0.0085</td>
<td>0.0282</td>
<td>0.104</td>
<td>0.2209</td>
<td>0.2898</td>
<td>0.3606</td>
<td>0.7177</td>
<td>0.7262</td>
<td>0.7731</td>
</tr>
<tr>
<td>90-8-10</td>
<td>0.0065</td>
<td>0.0066</td>
<td>0.0084</td>
<td>0.0282</td>
<td>0.1039</td>
<td>0.2208</td>
<td>0.2894</td>
<td>0.3598</td>
<td>0.7177</td>
<td>0.7262</td>
<td>0.773</td>
</tr>
<tr>
<td>85-10-10</td>
<td>0.0072</td>
<td>0.0072</td>
<td>0.0296</td>
<td>0.1047</td>
<td>0.2234</td>
<td>0.3078</td>
<td>0.3619</td>
<td>0.7191</td>
<td>0.7265</td>
<td>0.7776</td>
<td></td>
</tr>
<tr>
<td>90-5-10</td>
<td>0.0079</td>
<td>0.0081</td>
<td>0.0339</td>
<td>0.1177</td>
<td>0.2407</td>
<td>0.3113</td>
<td>0.367</td>
<td>0.7209</td>
<td>0.7264</td>
<td>0.8007</td>
<td></td>
</tr>
<tr>
<td>90-5-5</td>
<td>0.0078</td>
<td>0.0081</td>
<td>0.0379</td>
<td>0.1856</td>
<td>0.3113</td>
<td>0.2407</td>
<td>0.367</td>
<td>0.8007</td>
<td>0.7209</td>
<td>0.726</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 presents the results of this exercise for CAPM. The analysis of Figure 1 indicates that almost all the neural network models generated the best results up to 16 neurons while the combinations 60-20-20 and 90-05-05 dataset have ensured maximum performance in terms of mean squared error based on the CAPM-based market returns. A thorough search for the best network model in a data management software comes out with the 75-10-15 dataset and 16 neurons for the single factor CAPM as the best ANN model. The results of all other models started deteriorating after reaching the threshold number of the optimum neurons.

Figure 1: CAPM Based Results of All Dataset Combinations

The average monthly returns of the high, mid and low Beta portfolios are regressed on market return of CAPM. Figures 2 show the results of the fitness of CAPM based ANN system. Figure 2 indicates that the difference between the testing and validation error is subtle and shows an excellent fit of the model for high beta portfolios. This analysis refers to an important theoretical foundation of CAPM that the high-risk stocks have more chances to be predicted accurately with a neural network system. The results of high beta portfolios confirm the finding of Jabbari & Fathi (2014).
The results of the mid-Beta portfolio depicts widespread distance between testing and validation error. The wide gap between the testing and validation dataset shows an over-fitting of the network system. This overfitting in terms of the investment analysis means that the predictability of the mid beta portfolios has little chances of being accurate through a neural network system. This description of the mid beta portfolios is also confirmed by the regression analysis.

Figure 2: CAPM Based Performance for High, Mid and Low Beta Portfolios Respectively
The results of the low beta portfolios illustrate that the distance between testing and validation datasets is low as compared to the mid beta portfolios. These results are in line with the basic theory of the neural network system which state that ultimate error between the testing and validation dataset should be as small as possible. The observation of the time taken by the network is longer because the number of calculations is high as compared to the previous two systems. The R-value of the regression for the three classes of portfolios is described by Figure 3. These figures show that the regression line is lying at 45 degrees, which demonstrates a perfect relationship between the dependent and independent portfolio returns. The R-value for the high, mid and low beta portfolios is 99.99%, 90.52%, and 99.99% respectively. This also confirms the results of the graphs previously discussed and shows the correlation between actual and predicted portfolio returns for three classes of portfolios.

The findings suggest that the use of ANN in the capital asset pricing has successfully verified the theoretical foundation. The high beta portfolios have returned maximum predictability as compared to the mid and low beta portfolios. The corresponding sub-figures in Figures 3 also validate the theoretical postulation of CAPM. The regression score of the high beta portfolios are better than the mid and low beta portfolios. The regression value of high beta portfolios is almost 100% while the same score for mid and low beta portfolios is less than 100%.

A. Forecasting and Rolling Windows scheme in Asset Pricing

A rolling approach is applied to the sample data, which is designed for four years monthly returns. The neural networks take the values of monthly returns for the first 48 months as a training set and predict the 49th value for the monthly returns. The network then takes the next 48 values (leaving the first value and predicts the 50th value of monthly returns). This scheme of iterations generates a total of 133 values for MSE loss functions. It is pertinent to mention here that the program returns the best results only on an average basis (the average of training, validation, and testing). The rolling of the results on 48 months’ returns is the standard for the investment analysis in forecasting of the future values. The traditional forecasting theory suggests that the performance of forecasting techniques is evaluated on the out of sample data because the in-sample evaluation is of little use for the investors. Therefore, we have designed this new program wherein it returns the best results for the in-sample and out of sample (training, validation, and testing) separately for the loss function of MSE. This rolling method is adopted from Guerard et al. (2015).

B. Actual vs. Predicted Returns of Portfolio of CAPM based ANN

The actual and predicted portfolio returns of high, mid, and low beta portfolios are presented in Table 2. We processed the three classes of portfolios separately to investigate the best performance of ANN on various datasets and the optimal number of neurons. The upper part of Table 2 shows the MSE scores for each year’s returns of the high beta portfolios. The actual annualized returns of the high beta portfolios are ranging from -11.00% to 30.09% and the corresponding range of the predicted values of portfolio returns is -12.07% to 29.10%. These results are obtained from the optimal dataset of 75-15-15 and 16 neurons. The mean squared error is ranging from 0.0037 to 0.0161. The neural networks have proved its learning ability by predicting the negative and positive signs of the portfolio returns.
Figure 3: CAPM based Regression Analysis for Different Beta Portfolios
ow beta portfolios are in the range of results are also in line with the findings of Assi and show a fat tail. Due to these reasons, the predicted values of low and mid beta portfoliosthe predicted values for high beta portfolios is small as compared to the low and mid beta Gokgoz & Sezgin & Alp (2014) and Jasic & Wood (2004) and 8 neurons. These results and findings are in line with the findings of 69.32%. The low beta portfolios converge the best MSE statistics at 70 35. The l 35. The m 2015 2014 2013 2012 2011 2010 2009 2008 2007 2006 2005 2015 2014 2013 2012 2011 2010 2009 2008 2007 2006 2005 Year Actual Returns % Predicted Returns % MSE Data Set Optimal Neurons Actual vs. Predicted Returns (High Beta Portfolios) 2005 30.09 30.08 0.01 70-15-15 16 70.56% 41.36 19 0.0126 60-20-20 35 2009 -54.44 -54.45 0.0141 70-15-15 16 2010 -15.95 -15.96 0.013 70-15-15 16 2011 -19.19 -19.20 0.0161 70-15-15 16 2012 -11 -11.01 0.0108 70-15-15 16 2013 12.03 12.01 0.0117 70-15-15 16 2014 20.35 20.34 0.0055 70-15-15 16 2015 21.04 21.03 0.0037 70-15-15 16 Actual vs. Predicted Returns (Mid Beta Portfolios) 2005 22.24 22.23 0.009 60-20-20 35 2006 -14.2 -14.20 0.0088 60-20-20 35 2007 29.95 29.93 0.011 60-20-20 35 2008 -51.85 -51.85 0.0068 60-20-20 35 2009 -25.04 -25.05 0.0126 60-20-20 35 2010 -15.67 -15.68 0.0163 60-20-20 35 2011 -35.03 -35.04 0.0133 60-20-20 35 2012 55.68 55.67 0.0135 60-20-20 35 2013 15.29 15.28 0.0061 60-20-20 35 2014 15.29 15.30 0.0053 60-20-20 35 2015 40.56 39.98 0.0058 60-20-20 35 Actual vs. Predicted Returns (Low Beta Portfolios) 2005 55.52 55.50 0.0119 70-20-10 8 2006 23.75 23.73 0.0103 70-20-10 8 2007 -18.05 -18.06 0.0064 70-20-10 8 2008 41.37 41.36 0.0081 70-20-10 8 2009 -49.03 -49.03 0.0069 70-20-10 8 2010 -28.37 -28.38 0.0102 70-20-10 8 2011 -22.17 -22.18 0.01 70-20-10 8 2012 -25.45 -25.46 0.0146 70-20-10 8 2013 70.56 70.54 0.0124 70-20-10 8 2014 12.09 12.08 0.0065 70-20-10 8 2015 30.53 30.52 0.0059 70-20-10 8 The actual and predicted annualized returns of the mid beta portfolios are in Table 2. The actual and predicted returns are ranging from -51.85% to 55.68% and -52.53% to 54.33%. The optimal dataset is 60-20-20 and the optimal number of neurons is 35. The low beta portfolios are in the range of -49.03% to 70.56% and -49.72% to 69.32%. The low beta portfolios converge the best MSE statistics at 70-20-10 data set and 8 neurons. These results and findings are in line with the findings of Fadlalla & Amani (2014) and Jasic & Wood (2004).

An important conclusion of these Tables is that the volatility in the actual and predicted values for high beta portfolios is small as compared to the low and mid beta portfolios. The actual returns of the mid beta and low beta portfolios are highly skewed and show a fat tail. Due to these reasons, the predicted values of low and mid beta portfolios have shown more volatility as compared to high beta portfolios. This conclusion is according to the established finance practices which say that the market assigns higher compensation to high risky stocks as compared to low-risk portfolios. The results are also in line with the findings of Gokgoz & Sezgin-Alp (2014).
C. In-Sample and Out of Sample Analysis of CAPM Based Results

The MSE results for training, validation, and testing for all portfolios are shown in Table 3. The analysis of these results indicates that the MSE score for high beta portfolios is high on the initial rolling but gradually decreases for the training or in sample dataset, and the model has successfully reduced the testing or out of sample error. The network’s learning capability has steadily increased, and the model has not been trapped into local minima. The MSE score for the training set is ranging between 0.15% - 1.01%.

This enables us to conclude safely that the neural network system can efficiently train the stock market returns of a volatile market. The performance of the system on test data is even lower than training and validation datasets and MSE score for the experimental Dataset is ranging between 0.04% - 0.17%. These results verify the basic theory of ANN which says that the performance of the neural network systems increases when there is high nonlinearity in the dataset. Pakistan’s equity market is considered as the most volatile capital market, and the results lead to the conclusion that the application of ANN in this market can successfully capture the high nonlinearity trends in this market which can ultimately beat the passive buy and hold strategy. Table 3 shows that the mid beta and low beta portfolios have shown more variation regarding the forecasting error, but the model has performed in the same way as it did for the high beta portfolios.

Table 3 shows that the forecasted error for mid beta portfolios on the training set is ranging between 0.35% to 1.87% and the testing error is ranging between 0.04% to 0.05 %. While the low beta portfolios generated MSE score of 0.12% to 0.96% for training and 0.05 % to 0.62%.

Table 3: In-Sample and Out of Sample Results of CAPM (All Portfolios)

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>MSE Training</th>
<th>MSE Validation</th>
<th>MSE Testing</th>
<th>Sr. No</th>
<th>MSE Training</th>
<th>MSE Validation</th>
<th>MSE Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0.0022</td>
<td>0.0029</td>
<td>0.0008</td>
<td>P16</td>
<td>0.0187</td>
<td>0.0165</td>
<td>0.0005</td>
</tr>
<tr>
<td>P2</td>
<td>0.0108</td>
<td>0.0051</td>
<td>0.001</td>
<td>P17</td>
<td>0.0062</td>
<td>0.0029</td>
<td>0.0016</td>
</tr>
<tr>
<td>P3</td>
<td>0.0101</td>
<td>0.0017</td>
<td>0.001</td>
<td>P18</td>
<td>0.0095</td>
<td>0.0135</td>
<td>0.0013</td>
</tr>
<tr>
<td>P4</td>
<td>0.0043</td>
<td>0.0045</td>
<td>0.0014</td>
<td>P19</td>
<td>0.006</td>
<td>0.0028</td>
<td>0.0015</td>
</tr>
<tr>
<td>P5</td>
<td>0.0025</td>
<td>0.0026</td>
<td>0.0012</td>
<td>P20</td>
<td>0.0119</td>
<td>0.0048</td>
<td>0.0011</td>
</tr>
<tr>
<td>P6</td>
<td>0.0015</td>
<td>0.0022</td>
<td>0.0017</td>
<td>P21</td>
<td>0.0093</td>
<td>0.0041</td>
<td>0.0011</td>
</tr>
<tr>
<td>P7</td>
<td>0.003</td>
<td>0.0017</td>
<td>0.0004</td>
<td>P22</td>
<td>0.0096</td>
<td>0.0266</td>
<td>0.002</td>
</tr>
<tr>
<td>P8</td>
<td>0.0041</td>
<td>0.0024</td>
<td>0.0011</td>
<td>P23</td>
<td>0.0044</td>
<td>0.0072</td>
<td>0.0062</td>
</tr>
<tr>
<td>P9</td>
<td>0.0033</td>
<td>0.0031</td>
<td>0.0005</td>
<td>P24</td>
<td>0.0058</td>
<td>0.0094</td>
<td>0.0011</td>
</tr>
<tr>
<td>P10</td>
<td>0.0028</td>
<td>0.0019</td>
<td>0.0009</td>
<td>P25</td>
<td>0.0046</td>
<td>0.0018</td>
<td>0.0012</td>
</tr>
<tr>
<td>P11</td>
<td>0.0052</td>
<td>0.0062</td>
<td>0.0009</td>
<td>P26</td>
<td>0.0016</td>
<td>0.0084</td>
<td>0.0019</td>
</tr>
<tr>
<td>P12</td>
<td>0.0035</td>
<td>0.0031</td>
<td>0.0021</td>
<td>P27</td>
<td>0.0051</td>
<td>0.0077</td>
<td>0.0016</td>
</tr>
<tr>
<td>P13</td>
<td>0.0187</td>
<td>0.0018</td>
<td>0.0004</td>
<td>P28</td>
<td>0.0114</td>
<td>0.006</td>
<td>0.0012</td>
</tr>
<tr>
<td>P14</td>
<td>0.006</td>
<td>0.0067</td>
<td>0.0011</td>
<td>P29</td>
<td>0.0012</td>
<td>0.011</td>
<td>0.0022</td>
</tr>
<tr>
<td>P15</td>
<td>0.0103</td>
<td>0.0104</td>
<td>0.0012</td>
<td>P30</td>
<td>0.0069</td>
<td>0.0109</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

The corresponding Figures 4 of the three categories of portfolios confirm that the training curve shows an upward trend initially but gradually decreases which indicate that the network designed for this analysis has enabled the system to learn the patterns of the data. Once it is skilled in its learning process, it can generalize its parameters to any dataset. These findings are according to the theory of the neural network system which
says that the learning behavior of the system is a reflection of the human training, and it can adapt itself to the changing circumstances as humans behave.

**Figure 4:** In-Sample and Out of Sample Results for High, Mid and Low Beta Portfolios

| [Graph showing Mean Squared Error (MSE) for Training, Validation, and Testing for High Beta Portfolios] |
| [Graph showing Mean Squared Error (MSE) for Training and Validation for Mid Beta Portfolios] |
| [Graph showing Mean Squared Error (MSE) for Training and Validation for Low Beta Portfolios] |

**V. Conclusion**

The paper investigates the predictive ability of artificial neural networks by using the capital asset pricing model in Pakistan’s equity market. The major aim is to identify the ANN architecture which provides the best description of future returns. The principal measure is conducted through ANN and a total of thirty portfolios are constructed and categorized as high, mid and low beta portfolios. The initial testing of ANN is conducted through a wide range of parameters, while a rolling scheme is applied to the portfolio returns under the proposed method to estimate the time series of portfolio returns. This time series is predicted with the help of regression testing, curve fitness, the difference between the actual and predicted values of annualized returns and the difference between the training (in sample) and testing (out of sample). The study finds that the CAPM-based neural network models have predicted future returns accurately and the resultant errors of all the neural network models are minimizing in the range of 1-16 neurons. The goodness of fit test and the regression of the CAPM-based variable in the neural network system find a very close fit for the high and low beta portfolios but report a deviation for the mid beta portfolios. The regression values of the three classes of portfolios are approximately 100%, 90% and 100% respectively.
Under the rolling scheme of portfolio returns estimation, the MSE results for high, mid and low beta portfolios are 0.0015, 0.0024, and 0.0037. These results suggest that the CAPM-based ANN system has the capability to predict the future portfolio returns accurately. The predicted annualized returns of the high beta portfolios report very close results to the actual returns while the actual returns of mid beta and low beta portfolios are highly skewed and show a fat tail. Due to this reason, the predicted returns of these portfolios show high volatility as compared to high beta portfolios. This finding suggests that the market assigns higher compensation to risky assets. The high beta portfolios generate excess returns for the investors under the CAPM. These findings are in line with the findings of Fadlalla & Amani (2014) and Jasic & Wood (2004).

The results are useful for investors, especially in the stock market, who are interest in accurate forecasts. This will lead to better portfolio construction and optimal decision making. Furthermore, the investors can pursue passive portfolio decision making. Resultantly, the markets will represent less myopic behavior and move towards the stability of financial markets. The future researchers should work on the classification and clustering features of ANN in asset pricing and particle swarm optimization in the portfolio construction stage of ANN. A comparison of these results with the traditional forecasting techniques is another important area to explore.

References


