A Data Driven Model for Predicting the Financial Market Prices for Investors with Non Financial Experts: The Case of Saudi Arabia

Abdulaziz AlDaarmi
School of Engineering and Design
Brunel University
London, UK
abdulaziz.aldaarmi@brunel.ac.uk

Maysam Abbod
School of Engineering and Design
Brunel University
London, UK
Maysam.abbod@brunel.ac.uk

Abstract— the Fama and French (FF) three factors model has been one of most famous models in explaining the returns of stocks. This study investigates if the three factors of the FF model at the micro level may impact the return of portfolios in Saudi Arabia Stock Exchange (SASE) using an accurate modern technique in forecasting: Artificial Neural Networks (ANN). This study examined monthly data relating to common stocks from the listed companies of Saudi Arabia Stock Exchange from January 2007 to December 2011. The results from this study indicate that the FF model is applicable to measure the Stocks and portfolios reruns. Also, ANN technique can be used in predicting the stock and portfolios returns

Keywords: Fama and French (FF) Model, artificial neural networks (ANN) and stock exchange prediction

I. INTRODUCTION

Forecasting stock exchange price is a significant financial issue which has been receiving increased attention in the last few years. Different techniques are being used in the trading community for Forecasting stock exchange prices. In recent years, the new concept of the neural networks has emerged. One of these, the Artificial Neural Network (ANN), is able to create forecasts with significant predictive ability. ANN has been successfully applied in a variety of business fields including accounting [1], economics [2], finance [3], management information systems [4], marketing [5] and production management [6].

In this context, the aim of this study is to predict the stock price and determine whether the predictive power of stock price can be improved in Saudi Arabia Stock Exchange (SASE) by using the Artificial Neural Networks technique (ANN).

Moreover, this study explores the efficiency of Saudi Arabia Stock Exchange. If the attempts to improve the predicting power of stock price in (SASE) using the ANN technique made the market inefficient, then there are two possibilities. This inefficiency may be due to the fact that it is an emerging market. Alternatively, the predicing power of stock return in (SASE) cannot be improved by using this specific technique (ANN).

In order to achieve the aims of this study, the following objectives were set: (1) Determining the accuracy of computer-based information systems in predicting stock price movement for companies traded in SASE; (2) Specifying a model that may predict the stock return in SASE by applying the Fama and French (FF) model three factors at the micro level by using AAN.

This paper consists of five sections. Section 2 describes the literature review, while section 3 shows methodology of Fama and French. Section four lists the results, and finally the last section presents the conclusion.

I. LITERATURE REVIEW

This section presents the literature review in prediction stock exchange. Kim and Han used a ANN modified by Genetic Algorithm (GA) to predict the stock price index [7]. They concluded that the GA approach outperformed the conventional models. Kuo, Chen, and Hwang used a genetic algorithm base fuzzy neural network in the Taiwan stock market to measure the qualitative effects on the stock price [8]. Quek and Cheng used ANFIS and Neuro-Fuzzy network for forecasting investors in the US Stock Exchange [9], the findings indicate that ANFIS is effective for predicting stock prices in the US Stock Exchange. Abraham, Ramos, and Han used a genetic programming technique based on Multi-Expression programming (MEP) in order to forecast for two stock indexes [10], the results indicate that MEP is a novel and promising technique for function approximation problems. MEP technique gives the lowest MAP values for Nasdaq-100 index of Nasdaq Stock Market SM and the S&P CNX NIFTY stock index. Trinkle tested ANFIS and neural network to forecast the annual excess returns of the three publicly traded companies [11]. The predictive ability of these two techniques is compared with an autoregressive moving average (ARMA) model, the results show that the ANFIS and neural network techniques are able to create forecasts with significant predictive ability. Afolabi and Olatoyosiuse used Kohonen Self Organising Map (SOM) and hybrid Kohonen SOM for forecasting stock price, the results explain that the hybrid Kohonen self-organizing map is power predation than the other techniques [12]. Chang and Liu improved a Takagi–Sugeno–Kang (TSK) type fuzzy rule-based system for forecasting Taiwan Stock Exchange (TSE) price [13], the TSK fuzzy model efficiently forecasts stock price with accuracy close to 97.6% in TSE index and 98.08% in Media Tek. Abbasi and Aboueck investigated the current movement of stock price of the Iran Khodro Corporation at Tehran
Stock Exchange by using an Adaptive Neuro-Fuzzy Inference System (ANFIS) [14], the findings of the research demonstrate that the movement of stock price can be forecast with a low level of error. Guresen and Kayakutlu investigated neural network models like GARCH-DAN2 and EGARCH-DAN2, and compared these models in two regards: MSE and MAD to forecast Istanbul Stock Exchange (ISE XU10) [15]. Yunus, Shamsuddin and Sallehuddin developed a hybrid Neuro-Fuzzy with ANFIS to predict daily price of the Kuala Lumpur Composite Index (KLCI) [16], the results show that ANFIS method is competent in forecasting the KLCI compared to ANN. Atsalakis and Valavanis applied Adaptive Neuro-Fuzzy Inference System (ANFIS) to find the stock price prediction model. The results show that ANFIS are able to prediction the next day price of stock exchange [17].

II. METHODOLOGY

A. Introduction

In this study I will investigate a model that can predict the stock returns in SASE by applying the Fama and French (FF) three factors model using ANN. The three factors of the FF model are the market return, size and book-to-market ratio [18][19].

B. Data Description

The period of this study extended from January 2007 to December 2011, using monthly stock prices for corporations listed in Saudi Arabia Stock Exchange (SASE). The source of all the data used in this study is the website of the Saudi Arabia Stock Exchange. Therefore, the number of observations was 60 [20].

C. Monthly Return

The monthly return calculated in the following equation:

\[ R_{i,t} = (P_{i,t} - P_{i,t-1}) / P_{i,t-1} \]  

where \( R_{i,t} \) is the monthly return for a stock. 
\( P_{i,t} \) is the end of the month stock price 
\( P_{i,t-1} \) is the end of the previous month stock price

D. Forming the Dependent Variables Portfolios

This study covers all the companies in Saudi Arabia Stock Exchange; therefore, I divide the exchange companies upon 50% breakpoint for size at each year into two size groups: B for Big & S for Small. Then, each size group id divided into three book-to-market group upon two breakpoints 30% and 70% at each year. After that, six size portfolios (B/H, B/M, B/L, S/H, S/M, S/L) are formed from the two size portfolios. Table I below illustrates the six portfolios performed [21][22].

<table>
<thead>
<tr>
<th>SIZE</th>
<th>BOOK TO MARKET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above 70%</td>
<td>Between 70%- 30%</td>
</tr>
<tr>
<td>Above 50%</td>
<td>Big / High</td>
</tr>
<tr>
<td>Below 50%</td>
<td>Small / High</td>
</tr>
</tbody>
</table>

E. Forming the Independent Factors Portfolios:

From the six portfolios formed above in the dependent variable we construct the independent variables \( R_{SMB} \) (small minus big) portfolio returns which is defined as \( R_{SMB} = (R_{SL} + R_{SM} + R_{SH} - R_{SL} - R_{BM} - R_{BH})/3 \), and the HML (high minus low) portfolio returns which are defined as \( R_{HML} = (R_{BH} - R_{BL} - R_{SH} - R_{SL} - R_{BH} - R_{BL})/2 \). Also a value –weighted portfolio Market is formed which contains all the firms in these portfolios [21][22].

F. Equations

The equations of the three factors model of Fama and French (Fama and French [21]) are:

\[ R_{it} - R_{f} = \alpha_i + \beta_i (R_{M} - R_{f}) + \gamma_i S_{MB} + \delta_i H_{ML} + \epsilon_i \]  

The dependent variable is \( R_{i,t} - R_{f} \); the weighted average return for all the companies in stock market for six portfolio which are the following: (1) RHB, which is Portfolio return for companies that are high Book-to-Market level and big group; (2) RHS, which is Portfolio return for companies that are high Book-to-Market level and small group; (3)RMS, which is Portfolio return for companies that are medium Book-to-Market level and big group; (4) RMB, which is Portfolio return for companies that are medium Book-to-Market level and small group; (5) RLS, which is Portfolio return for companies that are low Book-to-Market level and big group and finally (6) RLS, which is Portfolio return for companies that are low Book-to-Market level and small group. The independent variables include the following. (1) \( R_{m} \): the market return portfolio is a sum over or aggregate portfolio of all individual investors, lending and borrowing will cancel out. In other words, it equals the entire wealth of the state economy [23]. The methodology of Fama and French [21] for \( R_{m} - R_{f} \) is the weighted average return of all the stocks in the sample. (2) \( R_{SMB} \): one of first and famous anomalies was size effect, which emphasizes that small size stocks had higher risk adjusted return than the stocks of the big size stocks [24]. The methodology of Fama and French [21] for \( R_{SMB} \) is explained by the difference between the return portfolios of small and big of stocks, by this equation: \( R_{SMB} = (R_{SL} + R_{SM} + R_{SH} - R_{BL} - R_{BM} - R_{BH})/3 \). (3) \( R_{HML} \): another famous anomaly was book-to-Market effect, which emphasizes that low market value stocks had poor prospects and must be penalized by higher risk adjusted return. [24].

The methodology of Fama and French [21], for \( R_{HML} \) is explained by the difference between the return on the portfolios of high and low-book-to-market stocks, through this equation: \( R_{HML} = (R_{BH} - R_{BL} - R_{SH} - R_{SL} - R_{BH} - R_{BL})/2 \).
III. RESULTS

The FF model three factors have been applied using Matlab software after normalizing the data, using the following techniques: (1) Logistic regression method (LR). (2) Different Neural Networks (NNs) types; namely, Feed Forward Network (NEWFF), Elman networks (NEWELM), Cascaded Forward Back Propagation (NEWCF), Radial Basis Networks (NEWRB), a feed-forward input time-delay back propagation network (NEWFFTD), a distributed time delay neural network (NEWDTDNN) and a fitting network (NEWFIT) [25][26][27][28][29]. The ANN parameters and topology are illustrated in Table II. (3) An Ensemble of the above Neural Network techniques. (4) An ensembled Neuro-Fuzzy (NF) system with 30 models, each having a different number of membership functions [30]. (5) Finally, all the ensembles were combined together using average weighting.

There were sixty monthly observations for each portfolio, divided into two parts. The first one contains the first 48 observations, which represent the training period from 2007 to 2010, while the last 12 observations (twelve months in 2011) represent the test period.

Tables III and IV show that the best result for standard deviation was Average Weighted Method (the average of the nine above portfolios) for all portfolios training and testing. Standard Deviation represents the difference between the actual values and the predicted ones. Figures 2, 4, 6, and 9 (RHB, RHS, RMB training portfolio and RLB testing portfolio) illustrate that the prediction power was very weak because the points were too far from the prediction line, which represents the predicted values or theoretical values. This means that the prediction accuracy is weak. On the other hand, Figures 1, 3, 7, 8, 10, 11 and 13 (RHB, RHS, RMB, RLS and RMS training portfolio and RMS, RLB, RLS testing) illustrate that the prediction power was very strong because the points were near and around the prediction line, which represents the predicted values or theoretical values. Consequently, the prediction accuracy is strong.

The FF three factors model can be adapted by investors in Saudi Arabia Stock Exchange because it makes high level of accrue predicting for stock prices in eight portfolios of the twelve portfolios. This means that we can depend on FF three factors model in a developing market like the Saudi Arabia Stock Exchange (SASE) for the prediction of stock price.

<table>
<thead>
<tr>
<th>TYPE</th>
<th>Topology</th>
<th>Train/valid</th>
<th>Training epochs</th>
<th>Training function</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEWFF</td>
<td>3-5-1</td>
<td>80/20</td>
<td>500</td>
<td>Levenberg-Marquardt</td>
</tr>
<tr>
<td>NEW ELM</td>
<td>3-5-1</td>
<td>80/20</td>
<td>500</td>
<td>Gradient descent</td>
</tr>
<tr>
<td>NEWCF</td>
<td>3-5-1</td>
<td>80/20</td>
<td>500</td>
<td>Levenberg-Marquardt</td>
</tr>
<tr>
<td>NEWRB</td>
<td>3-5-1</td>
<td>80/20</td>
<td>500</td>
<td>Radial Bases Functions</td>
</tr>
<tr>
<td>NEWFFTD</td>
<td>3-5-1</td>
<td>80/20</td>
<td>500</td>
<td>Levenberg-Marquardt</td>
</tr>
<tr>
<td>NEWDTDNN</td>
<td>3-5-1</td>
<td>80/20</td>
<td>500</td>
<td>Levenberg-Marquardt</td>
</tr>
<tr>
<td>NEWFIT</td>
<td>3-5-1</td>
<td>80/20</td>
<td>500</td>
<td>Levenberg-Marquardt</td>
</tr>
</tbody>
</table>

| TABLE III.  | Training Results for the Different ANNs and Average Weight |
| RAM:Training| rhb        | rhs | rms | rms | rlb | rhs |
| LR          | 0.25       | 0.22 | 0.19 | 0.18 | 0.20 | 0.22 |
| NEWFF        | 0.19       | 0.18 | 0.14 | 0.12 | 0.15 | 0.17 |
| NEWELM       | 0.27       | 0.29 | 0.23 | 0.27 | 0.26 | 0.26 |
| NEWCF        | 0.19       | 0.16 | 0.13 | 0.12 | 0.14 | 0.17 |
| NF-AVERAGE   | 8.3E-6     | 6.8E-6 | 8.2E-6 | 6.9E-6 | 4.5E-6 | 5.8E-6 |
| NEWRB        | 0.26       | 0.26 | 0.21 | 0.21 | 0.23 | 0.24 |
| NEWFFTD      | 0.21       | 0.19 | 0.17 | 0.13 | 0.14 | 0.17 |
| NEWDTDNN     | 0.20       | 0.19 | 0.15 | 0.13 | 0.13 | 0.19 |
| NEWFIT       | 0.21       | 0.19 | 0.15 | 0.12 | 0.14 | 0.17 |
| AVE-WEIGHT   | 0.18       | 0.17 | 0.13 | 0.12 | 0.13 | 0.16 |

| TABLE IV.   | Testing Results for the Different ANNs and Average Weight |
| RAM: Testing| rhb        | rhs | rms | rms | rlb | rhs |
| LR          | 0.10       | 0.36 | 0.34 | 0.08 | 0.22 | 0.27 |
| NEWFF        | 0.16       | 0.33 | 0.32 | 0.11 | 0.19 | 0.29 |
| NEWELM       | 0.15       | 0.26 | 0.26 | 0.15 | 0.19 | 0.19 |
| NEWCF        | 0.17       | 0.33 | 0.35 | 0.10 | 0.31 | 0.28 |
| NF-AVERAGE   | 0.24       | 0.14 | 0.30 | 0.24 | 0.30 | 0.24 |
| NEWRB        | 0.22       | 0.35 | 0.34 | 0.22 | 0.12 | 0.31 |
| NEWFFTD      | 0.21       | 0.35 | 0.30 | 0.08 | 0.24 | 0.24 |
| NEWDTDNN     | 0.10       | 0.33 | 0.33 | 0.09 | 0.18 | 0.23 |
| NEWFIT       | 0.20       | 0.35 | 0.34 | 0.11 | 0.19 | 0.24 |
| AVE-WEIGHT   | 0.15       | 0.32 | 0.30 | 0.10 | 0.17 | 0.22 |
Figure 1: RHB training results using average weight technique.

Figure 2: RHB testing results using average weight technique.

Figure 3: RHS training results using average weight technique.

Figure 4: RHS testing results using average weight technique.

Figure 5: RMB training results using average weight technique.

Figure 6: RMB testing results using average weight technique.
Figure 7: RMS training results using average weight technique.

Figure 8: RMS testing results using average weight technique.

Figure 9: RLB training results using average weight technique.

Figure 10: RLB testing results using average weight technique.

Figure 11: RLS training results using average weight technique.

Figure 12: RLS testing results using average weight technique.
IV. CONCLUSIONS
Predicting stock market price is a hard decision, but Artificial Neural Network (ANN) provides the ability to forecast stock price. This is a new and emerging area; there is a considerably large domain to use Artificial Neural Network (ANN) for predicting more accurate stock price as well as predict whether it is best to buy, hold or sell shares of stock market.

This paper investigated the FF model three factors using ANN to prediction the movement of stock price in the future in Saudi Arabia Stock Exchange (SASE). The results show that the ANN techniques are able to generate forecasts with significant predictive ability. The ANN can improve the investor’s prediction for the stock price in SASE.

ACKNOWLEDGMENTS
The author wishes to thank the ministry of Higher Education in Saudi Arabia for providing the financial support to do this project.

REFERENCE


http://www.tadawul.com.sa/