Advances in mathematical finance & applications, 4 (3), (2019), 1-13
DOI: 10.22034/amfa.2019.584494.1173

Published by IA University of Arak, Iran
Homepage: www.amfa.iau-arak.ac.ir

Improving Stock Return Forecasting by Deep Learning Algorithm

Zahra Farshadfar [a, *], ***Marcel Prokopczuk*** [b]

[a]*Department of Economics, College of Humanities, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran.*
[b]*Institute for Financial Markets, Leibniz University Hannover, Hannover, Germany.*

ARTICLE INFO

Article history:
Received 3 February 2019
Accepted 30 May 2019

Keywords:
Deep Learning
Historical Average Model
Nonlinear Model
Gold Price

ABSTRACT

Improving return forecasting is very important for both investors and researchers in financial markets. In this study we try to aim this object by two new methods. First, instead of using traditional variable, gold prices have been used as predictor and compare the results with Goyal's variables. Second, unlike previous researches new machine learning algorithm called Deep learning (DP) has been used to improve return forecasting and then compare the results with historical average methods as bench mark model and use Diebold and Mariano's and West's statistic (DMW) for statistical evaluation. Results indicate that the applied DP model has higher accuracy compared to historical average model. It also indicates that out of sample prediction improvement does not always depend on high input variables numbers. On the other hand when using gold price as input variables, it is possible to improve this forecasting capability. Result also indicate that gold price has better accuracy than Goyal's variable to predicting out of sample return.

1 Introduction

Research on the predictability of stock markets has a long history in financial economics [7]. While opinions differ on the efficiency of markets, many widely accepted empirical studies show that financial markets are to some extent predictable [6]. The main point is when the goal is to improving return prediction, most researches are trying to overcome this problem by increasing the variable. Less research has been focus on changing the model to improve the prediction or if this happened, most of these models are linear and less are considered for nonlinear models. Despite the fact that researches show that so many factors are affecting the stock market's performance, so fluctuations in the stock market are non-linear [18]. Another point is, however, the literature finds relatively limited evidence of predictability using out-of-sample tests. The general conclusion is that the evidence for stock return predictability is predominantly in-sample; and it is not robust to out-of-sample evaluations [28]. In this study we focus on this main point. we want to see if out-of-sample prediction improvement does not only depend on increase in the number of input variables and besides that, can sometimes better result be obtained by choosing less but correct number of variables as well as changing the model. Then the robustness of forecasting performance has been evaluated by allocation of different forecast-

*** Corresponding author Tel.:** *+9809183893673*
E-mail address: zfarshadfar@iauksh.ac.ir

ing model at different data frequencies to see if the results depend only on the data frequency or also on the model. We use daily, weekly, and monthly S&P 500 indices over the period 4 January 1986 until 31 December 2017. Three forecasting approaches for the out-of-sample stock returns based on the gold price are utilised. They are the Ordinary Least Squares (OLS), artificial neural network (ANN) and developed of artificial neural network called 'Deep Learning Algorithm'. In-sample periods of 30%, 50% and 70% are chosen. This is important since the literature has shown that the results can vary depending on the in-sample periods, particularly in the finite sample sizes. The evaluation statistics for forecasting performance are out-of-sample R-squared (R^2_{OS}) and Diebold-Mariano test.

2 Literature Review

Forecasting in the financial time series is basically predicting the series behavior one or few steps ahead with the help of a number of variables. The variables used for forecasting are either economic[1] or financial variables[2] or in some cases, technical analysis output [13]. In some studies, these variables are used in combination [29, 32]. One of the benefit of variable combination is obtaining information from many economic and financial variables and at the same time forecast volatility reduction [28]. It has to be mentioned that even though financial and economic data are different in nature, their combination should be done with caution. This research employs the gold price as a predictor of stock returns. There are some reasons why gold price has been chosen as a predictor. First, gold is a strategic commodity that has recently received much attention because it is among the most popular economic indicators and its price has lately exhibited a high volatility and may affect the prices in the commodity markets. There are so many studies exist that use gold price to study its effects on stock returns. These literature reports some kinds of relation between gold price and stock returns. Gold as a leader in the precious metals, has been used as money through history. While currently it is used for industrial purposes, its demand mostly originates from investment and jewellery. With jewellery demand stable and industrial demand declining, demand has increased because of retail investment especially after 2006 [27].

In fact gold is a basic part of most of the countries' international reserves portfolio and part of most investors' portfolio around the world. In addition, investors buy gold to hedge the increasing risk in financial markets. Gold is one of the risk management tools used in hedging and diversifying commodity portfolios. Moreover, investors in both advanced and emerging markets often switch between oil and gold or combine them to diversify their portfolios [31]. Second there is a common belief that the prices of crude oil and gold, tend to move in unison. The reason why commodity prices tend to rise and fall together is because they are influenced by common macroeconomic factors such as exchange rates and inflation [17]. Generally, the relationship between oil price and gold price is positive, implying that oil and gold are close substitutes as safe havens from inflation. Theoretically and according to empirical studies, oil and gold prices are indirectly linked through inflation[3], revenue

1 Economic variable like: consumption-wealth ratio, inflation rate, nominal interest rates and price ratio.
2 Financial variables like: book-to-market ratio, corporate issuing activity, dividend, earnings-price ratio, term and default spreads.
3 Many researchers examined the inflation hedge effectiveness of gold by analyzing the short-run and long-run relationship between the inflation rate and the price of gold. Generally speaking, the evidence regarding a long-run relationship for both linearity and nonlinearities between gold and consumer prices strongly depends on the sample period and is not clear-cut although a reasonable conclusion is that gold is at least partially able to provide a hedge against inflation.

channel[4], and exchange rates[5]. According to macroeconomic theory, the relationship between the two commodities could stem from the fact that oil price shocks may have a negative impact on the real economy and share prices; leading investors to look for alternative assets such as gold in terms of a hedge or a safe haven function. To sum up, the preceding literature shows that these two commodities are indirectly linked through different channels. However, fundamentally speaking, the direct relationship between oil and gold prices has not yet been considered. Empirically, some of the existing studies provided evidence that oil prices impact gold prices [24, 31].

It has been suggested for a long time that return can be predicted using different kind of methods mainly classified as time series forecasting models. These model can be divided in two big categories: classical models or econometrical models, that more than of them are liner, and modern model that most of them are nonlinear and usually originate from other disciplines such as mathematics, physics, computer and etc [19, 20, 30, 26, 14]. Empirical studies result on return forecasting models are different. Some of them propose nonlinear models do not perform linear models [21, 22], some others show linear models outperform or perform as well as nonlinear models [1, 2], and finally some studies find nonlinear models outperform linear models [12, 30].

Although stock market is a non-linear dynamic system and predicting the stock prices path is a difficult task, but there is uniform agreement that stock returns behave nonlinearly [9, 3] and if we show this nonlinear behavior linearly, the model has not been chosen correctly and we encounter model specification Error problem. Artificial Neural Network is one of the nonlinear return forecasting models that have virtually no theoretical inspiration in finance (they are often labeled black box technology), but it owes its popularity to its ability to fit into any functional data relationship to an arbitrary degree of exactness and works best in circumstances where financial theories have virtually nothing to say about likely functional form for the relationship between a set of variables [5].In addition, this model can discover the complex nonlinear relations and handle the predominant uncertainty and inaccuracy in the stock market [3, 11].Despite its abilities, this model has not been very privileged among economists. Studies have shown that this model has not succeeded in predicting out-of-sample data. Therefore, in order to obtain a better result, this model is usually combined with other nonlinear models like fuzzy family model or genetic algorithm, etc. This combination improves the model's performance and at the same time, it increases the complexity of the method too.

From the many available kind of intelligence algorithm ‘Deep Learning Algorithm’ has been chosen. There are some reason for this choice: initially, deep learning is one of the machine learning algorithms which has received considerable attention in recent years. There seems to be increasing interest in whether deep learning can be efficiently applied to financial problems. But the literature (at least in the public domain) still remains limited. Second: classical ANN models have structural problems such as gradient vanishing that doesn’t allowed the researcher to increase the layer or node and random weighting to the first layer parameter and vice versa. These problems have limited the forecasting ability of the model. The mentioned have been solved in deep learning method. In fact, this method offers an innovative technique that does not necessitate a pre-specification during the modeling pro-

4 Oil and gold prices are also indirectly linked through revenue channels. Some researches point out that gold price could also change with increasing oil prices if oil-exporting countries invest their revenues in gold. In addition, when oil price increases, oil exporters’ revenues from oil rise as well. If gold represents a significant share of the asset portfolio of oil exporters, and oil exporters purchase gold in proportion to their asset portfolio, then the gold price rises. In that case, an oil price rise leads to a rise in gold price

5 Some researchers argue that if a commodity such as gold or oil is denominated in dollars, a domestic appreciation against the dollar lowers the price of oil or gold in terms of the domestic currency, which increases demand and may lead to an increase in oil or gold prices.

cess as they independently learn the relationship characteristic in the variables [9]. Finally, every day many data are generated in financial markets that affects investors' expectations, supply, demand, and ultimately prices. In classical models it is impossible or very difficult to study these data simultaneously and when it has been done, it is difficult to provide model presumptions as many of the data's information disappears. Deep learning is nonlinear data based method that takes huge amount of data as input and then extracts information from them. This is why this model is usually used in condition that using true preprocessing data methods can keep many of data related information. Our contributions to the literature are summarized as follows. First, we contribute further insights for the gold price and stock market relationship literature by employing an out-of-sample analysis utilizing a recently developed of Artificial neural network called Deep learning Algorithm to show how well the gold price forecasts stock returns. Second, we test the robustness of forecasting performance by using different forecasting model. Finally we compare the out of sample forecasting ability of gold price with some Goyal's variables.

3 Proposed Methodology

Previous study such as Rapach *et al* [29] and Westerlund and Narayan [33] have tested the stock return predictability based on the simple predictive regression model as follow:

$$r_{t+h} = \alpha + \beta x_t + \epsilon_{t+h} \tag{1}$$

where r_{t+h} is the excess stock return, measured as stock returns in excess of the risk-free rate and x_t is the predictor variable (in this study gold price).

The predictive ability of x_t can be tested under the null and alternative hypothesis as follow:

$$\begin{cases} H_0{:}\beta=0 \ \rightarrow \text{no predictability} \\ H_1{:}\beta\neq 0 \rightarrow \text{xt does predict future stock returns.} \end{cases} \tag{2}$$

Although this model is simple to estimate, Phan *et al.* [28] stated that this model suffers from endogeneity, persistency, and heteroscedasticity issues. This while, the mentioned problem have been solved for machine learning models[6].

In this research we have use this predictive regression as follow: first, we have use it for estimating traditional linear model, OLS to compare it with nonlinear models. Second, search a predictor function f in order to predict the market return at time t+1 , r_{t+1} , given the features (representation) u_t extracted from the information available at time t. We assume that r_{t+1} can be disjointed into two parts: the predicted output $\hat{r}_{t+1} = f(u_t)$, and the unpredictable part γ, which we regard as Gaussian noise:

$$r_{t+1} = \hat{r}_{t+1} + \gamma, \qquad \gamma \sim N(0, \beta) \tag{3}$$

Where N (0, β) denotes a normal distribution with zero mean and variance β. The representation u_t can be either a linear or a non- linear transformation of the raw level information R_t. Denoting the transformation function by φ, we have:

6 Data preprocessing or data representation is the first level in machine learning algorithm and before model estimated this problem had been solved by feature engineering.

$u_t = \varphi(R_t)$ (4)

$\hat{r}_{t+1} = f \circ \varphi(R_t)$ (5)

When we use artificial neural network as nonlinear model, the nonlinear relationship between two variables h_l and h_{l+1} is specified by neural network through a network function, which typically has the form of:

$h_{l+1} = \delta(W h_l + b)$ (6)

In which, δ is called an activation function, and the matrix W and vector b are model parameters. The variables h_l and h_{l+1} are said to form a layer; when there is only one layer between the variables, their relationship is called a single-layer neural network.

Multi-layer neural networks augmented with advanced learning methods are generally referred to as deep neural networks (DNN). DNN for the predictor function, $y = f(u)$, can be constructed by serially stacking the network functions as follows:

$y = \delta_L (W_L h_{L-1} + b_L)$ (7)

Where L is the number of layers.

Given a dataset $\{u^n, \tau^n\}^N_{n=1}$ of inputs and targets, and an error function $\varepsilon(y^n, \tau^n)$ that measures the difference between the output $y^n = f(u^n)$ and the target τ^n, the model parameters for the entire network, $\theta = \{W_1,\dots, W_L, b_1,\dots,b_L\}$, can be chosen so as to minimize the sum of the errors:

$$\min_{0} \left[J = \sum_{n=1}^{N} \in (y^n, \tau^n) \right] \quad (8)$$

Given an appropriate choice of ε (·), its gradient can be obtained analytically through error back propagation (Bishop, 2006). In this case, the minimization problem in (8) can be solved by the usual gradient descent method. A typical choice for the objective function that we adopt in this paper has the form:

$$J = \frac{1}{N} \sum_{n=1}^{N} \|y^n - \tau^n\|^2 + \lambda. \sum_{l=1}^{L} \|W_1\|_2 \quad (9)$$

Where $\|.\|$ and $\|.\|_2$ respectively denote the Euclidean norm and the matrix L_2 norm. The second term is a “regularizer” added to avoid over fitting, while λ is a user-defined coefficient.

3.1 Training Deep Neural Networks

We constructed the predictor functions, $\hat{r}_{i,t+1} = f_i(u_t), i = 1, \dots, M$. Using DNN, we employ a three-layer network with 100 nodes in hidden layers of the form:

$$\begin{aligned} h_1 &= \mathrm{ReLU}(W_1 u_t + b_1) \\ h_2 &= \mathrm{ReLU}(W_2 h_1 + b_2) \\ \hat{r}_{i,t+1} &= W_3 h_2 + b_3 \end{aligned} \quad (10)$$

Where ReLU is the rectified linear unit activation function defined as: ReLU(x) = max (x, 0) with max being an element-wise operator. ReLU is known to provide a much faster learning speed than standard sigmoid units, while maintaining or even improving the performance when applied to deep neural networks [23]. Nevertheless, we also test a standard ANN for comparison: we replace ReLU with a sigmoid unit and reduce the nodes in hidden layers to one and keep the rest of the settings and parameters equal. With one sets of features, Raw Data, as the input, the DNN is trained minimizing the objective function defined in Eq. (9) with 1000 learning iterations (epochs), and the regularizer coefficient, λ=0.001. The last 20% of the training set is used as a validation set for early stop- ping to avoid over fitting. The training was conducted using Keras in Anaconda 3, Python 3.6.

3.2 Benchmark Model and Statistical Evaluation

Following previous studies such as Welch and Goyal [32] and Rapach *et al.* [29], we use historical average model as bench mark model:

$$\hat{r}_{T_0+1} = \frac{1}{T_0}\sum_{j=1}^{T_0} r_j \tag{11}$$

For Statistical evaluation according to Campbell and Thompson [8], R2 statistic has been chosen. It is computed as fallow:

$$R^2_{OS} = 1 - \frac{MSFE_1}{MSFE_0} \tag{12}$$

$MSFE_0$: MSFE of the benchmark model.

$MSFE_1$: MSFE of the competitor model.

The R^2_{OS}measures the decrease in MSFE for the opponent model compared to the benchmark model. If the opponent model's MSFE is less than that of the benchmark model (R^2_{OS}> 0), it shows that the opponent model is more accurate in forecasting than the benchmark model.

We also test the null hypothesis $H0$: $MSFE0 \leq MSFE1$ against $MSFE0 > MSFE1$, corresponding to $H0$: $ROS2 \leq 0$ against $ROS2 > 0$, the most common method incidentally is Diebold and Mariano's [10] and West [34] statistic (DMW), which has the next form:

$$DMW = \sqrt{(T - T_0 - h + 1)}\,\frac{\bar{d}}{\sqrt{\hat{s}}} \tag{13}$$

Where $\bar{d} = \frac{1}{(T - T_0 - h + 1)}\sum_{t=T_0}^{T} \hat{d}_{t+h}$ (14)

$$\hat{d}_{t+h} = \hat{u}^2_{0,t+h} - \hat{u}^2_{1,t+h} \tag{15}$$

$$\hat{u}_{0,t+h} = \hat{r}_{t+h} - \hat{r}_{0,t+h} \tag{16}$$

$$\hat{u}_{1,t+h} = \hat{r}_{t+h} - \hat{r}_{1,t+h} \tag{17}$$

$$\hat{s} = \frac{1}{T - T_0 - h + 1}\sum_{t=T_0}^{T}\left(\hat{d}_{t+h} - \bar{d}\right)^2 \tag{18}$$

4 Data Analysis and Findings

We constructed a Deep Neural Network by using total stock returns from monthly prices of the S&P500 index from 4 January 1986 until 31 December 2017. Stock returns were collected from the Bloomberg database, while the gold price obtained from World Gold Council. Stock returns are measured as a continuous compounded return $r_t = \ln(p_t/p_{t-1})$, and the three-month Treasury bill rate is used to calculate the excess returns. We use different in-sample periods with the proportions 30%, 50%, and 70% of the full sample to forecast the out-of-sample stock returns. As a result, the three out-of-sample periods are Jun 1995 to December 2012, November 2001 to December 2012, and April 2008 to December 2012.

We employ three different predictive model, namely, the OLS, ANN, and DNN to forecast the stock returns of the S&P500 index using the gold price as predictor. The out-of-sample forecasting accuracy was evaluated by the well-known $R^2{}_{OS}$ statistic. To test for the statistical significance of the outperformance, the DMW statistic was used for 4 model comparisons. The forecasting evaluation is undertaken for the four horizons.

4.1 Models Comparison with Historical Average (Gold Price as Predictor)

In this part we reported results which compared the OLS, ANN, and DNN predictive models based on gold price as a predictor with the historical average model as bench mark based on the monthly data frequencies. Starting with Table 1, in this table for 30%, 50% and 70% in sample period and h=1, 3, 6, 12 forecast horizon R^2_{OS} is negative. For example, for h=1 forecast horizon and 30%, 50% and 70% in sample period R^2_{OS} rang from -1.08 to -1.53 and for h=12 forecast horizon and same in sample period R^2_{OS} rang from -4.6 to -6.6.

Table 1 presents the results of the comparison between the OLS predictive regression models to the historical average model. The table reports the R^2_{OS}statistic and the p-value of the MSFE-adjusted statistic for assessing the statistical significance of the corresponding forecasts under the null hypothesis that the opponent forecasts (OLS) are not better than the benchmark forecasts (historical average).

Table 1: OLS versus historical average in stock return prediction.

Monthly Data								
GOLD	h=1		h=3		h=6		h=12	
30%	-1.148	(0.23)	-3.318	(0.52)	-4.650	(0.84)	-4.671	(0.35)
50%	-1.080	(0.15)	-4.34	(0.46)	-5.34	(0.52)	-5.45	(0.34)
70%	-1.530	(0.18)	-4.45	(0.63)	-5.62	(0.61)	-6.671	(0.54)

Source: Research Results.

Focusing on the OLS estimator results, since the R^2_{OS} statistics are negative in all cases, we found that the OLS estimator (with gold price as predictor) cannot worn out the historical average model. In addition, the results are consistent across the choices of in-sample periods and forecasting horizons. We conclude that the OLS based gold price estimators are not better than the historical average model in forecasting the stock returns. For the ANN model we observe similar result as OLS. Table 2 presents the results of the comparison between the ANN predictive models to the historical average model. In this table for 30%, 50% and 70% in sample period and h=1, 3, 6, 12 forecast horizon R^2_{OS} is negative.

For example, for h=1 forecast horizon and 30%, 50% and 70% in sample period R^2_{OS} rang from -0.027 to -0.082 and for h=12 forecast horizon and same in sample period R^2_{OS} rang from -0.0059 to -0.061. it appearances that the opponent model's MSFE (ANN) is greater than the benchmark model (R^2_{OS}<0), it shows that the opponent model is less accurate in forecasting than the benchmark model and the ANN model with gold price as predictors cannot beat the historical average model, , since R^2_{OS} statistics are negative in all cases. The results are consistent across the choices of in-sample periods and forecasting horizons.

Table 2: ANN versus historical average in stock return prediction.

Monthly Data								
GOLD	h=1		h=3		h=6		h=12	
30%	-0.053	(0.29)	-0.052	(0.32)	-0.050	(0.25)	-0.051	(0.27)
50%	-0.081	(0.27)	-0.082	(0.29)	-0.071	(0.29)	-0.061	(0.25)
70%	-0.027	(0.23)	-0.027	(0.31)	-0.024	(0.26)	-0.0059	(0.22)

Source: Research Results.

Then the DNN models with gold price as predictor was compared with the historical average model (results can be seen in Table 3). We observed that, the results are in favour of DNN models over the historical average model in forecasting stock returns in some in sample period. It was also observed that superiority of this model compared to historical average model is statistically significant in some forecasting horizon. Briefly, when monthly data were used for gold price, we obtain both positive and negative R^2_{OS}. The results based on monthly data frequencies strongly suggest that the DNN estimator is better than the historical average benchmark model, as the R^2_{OS}takes positive values in h=6 horizon of 50% in-sample periods (R^2_{OS}= 0/046) and 70% in-sample periods (R^2_{OS} = 0/039). For the 12 month forecast horizon and 70%, 50% and 30% in sample period for monthly data frequency R^2_{OS} range from 0/19 to 0/33 for the gold price and the results are better than the historical average model. For h=12 and 70% forecast horizon R^2_{OS} statistically significantly greater than zero.

Table 3: DNN versus historical average in stock return prediction.

Monthly Data								
GOLD	h=1		h=3		h=6		h=12	
30%	-0.06665	(0.23)	-1.9241	(0.31)	-0.0750	(0.61)	0.19740	(0.57)
50%	-0.04883	(0.45)	-1.6782	(0.25)	0.0468	(0.68)	0.33662	(0.048)**
70%	-0.02958	(0.56)	-1.9599	(0.61)	0.0397	(0.69)	0.31357	(0.71)

Source: Research Results. *, **, *** denote significance at the 10%, 5% and 1% levels.

4.2 Compare some Goyal's Variable with Gold Price as Predictor

In this part we focused on ANN and DNN predictive models based on inflation (inf), the term spread (tms)[7] and percent equity issuing (eqi)[8] .They are some of Welsh and Goyal [32] variables, that has

[7] The term spread is (tms) is the difference between the long term yield on government bonds and the treasury –bill, (Welch and Goyal, [32]).

[8] precent equity issuing (eqi) is the ratio of equity issuing activity as a fraction of total issuing activity, (Welch and Goyal, [32]).

shown good results in monthly frequencies out of sample return forecasting. These variable are: inflation (inf), the term spread (tms) and percent equity issuing (eqi). We compare them with gold price as a predictor with two machine learning algorithm ANN and DNN method.

Table 4: ANN versus historical average

Monthly Data								
GOLD	h=1		h=3		h=6		h=12	
30%	-0.053	(0.29)	-0.052	(0.32)	-0.050	(0.25)	-0.051	(0.27)
50%	-0.081	(0.27)	-0.082	(0.29)	-0.071	(0.29)	-0.061	(0.25)
70%	-0.027	(0.23)	-0.027	(0.31)	-0.024	(0.26)	-0.0059	(0.22)
TMS	h=1		h=3		h=6		h=12	
30%	-0.03598	(0.21)	-0.03539	(0.22)	-0.03356	(0.25)	-0.0345	(0.31)
50%	-0.01297	(0.19)	-0.01214	(0.20)	-0.00527	(0.23)	-0.00156	(0.29)
70%	0.0075	(0.18)	0.007451	(0.19)	0.0019	(0.20)	0.004161	(0.27)
EQI	h=1		h=3		h=6		h=12	
30%	-0.08865	(0.27)	-0.08772	(0.30)	-0.08492	(0.27)	-0.08709	(0.26)
50%	-0.03537	(0.29)	-0.03524	(0.29)	-0.03088	(0.31)	0.051606	(0.19)
70%	-0.0181	(0.24)	-0.02118	(0.28)	0.033745	(0.26)	0.032541	(0.21)
INF	h=1		h=3		h=6		h=12	
30%	-0.04683	(0.32)	-0.04707	(0.29)	-0.04552	(0.26)	-0.0444	(0.17)
50%	-0.00982	(0.31)	-0.01022	(0.19)	-0.01009	(0.29)	-0.01247	(0.26)
70%	-0.03349	(0.34)	-0.03764	(0.26)	-0.15413	(0.31)	-0.14284	(0.35)

Source: Research Results. *, **, *** denote significance at the 10%, 5% and 1% levels.

The data are collected from the Goyal site from 4 January 1986 to 31 December 2017. Starting with the ANN model and table number 4, we observe that all of the R^2_{OS}statistics are negative for all variable and in all forecast horizon except 70% in-sample period for all forecast horizons of tms variable (R^2_{OS} = 0.0075, 0.0074, 0.0019, 0.0041) and 70% in-sample period For 6 (R^2_{OS} = 0.033) and 12 (R^2_{OS} =0.032) month forecast horizons of eqi. Although these R^2_{OS} are positive but not statistically significant. Focusing on the DNN model and Table 5, when we used monthly data, for all variable, we obtain both positive and negative R^2_{OS}. Expect percent equity issuing and inflation. They have negative R^2_{OS} for all forecasting horizons and all in-sample period. For gold price the R^2_{OS}takes positive values in h=6 horizon of 50% in-sample periods (R^2_{OS}= 0/046) and 70% in-sample periods (R^2_{OS} = 0/039). For the 12 month forecast horizon and 70%, 50% and 30% in sample period for monthly data frequency R^2_{OS} from 0/19 to 0/33 for the gold price. For h=12 and 70% forecast horizon R^2_{OS} statistically significantly greater than zero.

For the term spread (tms) we obtain positive R^2_{OS} for all forecasting horizons and 70% in-sample period (R^2_{OS} =0.205 for h=1, R^2_{OS} =0.200 for h=3, R^2_{OS} =0.275 for h=6, R^2_{OS} =0.308 for h=12).

Table 5: DNN versus historical average

Monthly Data								
GOLD	h=1		h=3		h=6		h=12	
30%	-0.06665	(0.23)	-1.92418	(0.31)	-0.07503	(0.61)	0.197407	(0.57)
50%	-0.04883	(0.45)	-1.67823	(0.25)	0.04686	(0.68)	0.336622	(0.048)**
70%	-0.02958	(0.56)	-1.95997	(0.61)	0.0397	(0.69)	0.313577	(0.71)
TMS	h=1		h=3		h=6		h=12	
30%	-0.02781	(0.28)	-0.02744	(0.19)	-0.02642	(0.25)	-0.02715	(0.27)
50%	-0.02133	(0.24)	-0.02136	(0.22)	-0.0191	(0.29)	-0.01678	(0.34)
70%	0.20566	(0.21)	0.200057	(0.22)	0.275688	(0.31)	0.308491	(0.19)
EQI	h=1		h=3		h=6		h=12	
30%	-0.02632	(0.31)	-0.02631	(0.22)	-0.02669	(0.31)	-0.02604	(0.31)
50%	-0.03052	(0.32)	-0.0304	(0.41)	-0.03171	(0.19)	-0.03189	(0.25)
70%	-0.02135	(0.19)	-0.02141	(0.32)	-0.02207	(0.24)	-0.0203	(0.29)
INF	h=1		h=3		h=6		h=12	
30%	-0.07829	(0.51)	-0.07831	(0.38)	-0.07921	(0.33)	-0.07758	(0.45)
50%	-0.04026	(0.58)	-0.04011	(0.48)	-0.03968	(0.29)	-0.03861	(0.48)
70%	-0.0303	(0.61)	-0.03038	(0.36)	-0.03257	(0.21)	-0.03046	(0.51)

Source: Research Results. *, **, *** denote significance at the 10%, 5% and 1% levels.

5 Conclusions and Suggestions

This paper has used S&P500 indices gold price for stock returns forecast. It has used DNN forecasting model, which has solved some problems related to ANN model such as gradient vanishing and randomly weighted matrix. Monthly data over the period from 4 January 1986 to 31 December 2017 has been used. Three choices of in-sample periods with the proportions of 30%, 50% and 70% of the full sample were utilized for data mining and to overcome the over fitting in order to predict the out-of-sample stock returns. Finally, some of Welsh and Goyal [32] variables, that has shown good results in monthly frequencies out-of-sample prediction, are collected from the Goyal site from 4 January 1986 to 31 December 2017. These variable are: inflation (inf), the term spread (tms) and percent equity issuing (eqi). We compare them with gold price as a predictor with two machine learning algorithm ANN and DNN method.

The main findings of this research are: First, unlike the existing literature on gold price and stock returns, it has focused on out-of-sample prediction of returns. It has showed that the estimator has a great role and how well the gold price forecasts stock returns depend on the estimator. While this message is consistent with Westerlund and Narayan [33], Narayan and Sharma [25], phan and *et al.* [28] and Farshadfar and Prokopczuk [14], we add to this literature by showing that the relevance of the estimator matters not only when using traditional predictors of stock returns but also when using non-traditional stock return predictors, such as gold prices. Pair-wise assessments has been employed between the OLS, ANN and DNN models with historical average benchmark model. It has been observed that the DNN model using performs better than the historical average model in forecasting stock returns when the gold price is used as a predictor. This has been opposite for the OLS and ANN models. This finding is consistent with Chong *et al.*[9], Farshadfar and Prokopczuk[14].

The key suggestion of this finding is that:
First, the choice of the estimator in forecasting returns is important and deserves particular attention.
Second, while traditional financial ratios and selected macroeconomic variables have been found to predict stock returns, non-traditional variables which have serious economic repercussions such as gold price should not be ignored.
Third, in our study, gold price is not only statistically related to stock returns but there is a strong economic connection as well. This finding also points toward a need to explore other non-traditional predictors of returns amongst others.

Acknowledgment

Sabbatical financial support provided by Kermanshah branch of Islamic Azad University is appreciated.

References

[1] Adebiyi, A., Adewumi, A., Ayo, C., *Comparison of ARIMA and artificial neural networks models for stock price prediction*, Journal of Applied Mathematics, 2014. Doi: 10.1155/2014.614342.

[2] Agrawal, J., Chourasia, V., Mittra, A., *State-of-the-art in stock prediction techniques. International Journal of Advanced Research in Electrical*, Electronics and Instrumentation Engineering, 2013, **2**(4), P.1360–1366.

[3] Ahmadi, E., Jasemi, M., Monplaisir, L., Nabavi, M., Mahmoodi., A., Amini Jam, P., *New efficient hybrid candlestick technical analysis model for stock market timing on the basis of the support vector machine and heuristic algorithms of imperialist competition and genetic*, Expert Systems with Applications, 2018, **94**(1), P.21-31. Doi:10.1016/j.eswa.2017.10.023.

[4] Bishop, C., M., *Pattern Recognition and machine learning*, Springer, Germany, 2006.

[5] Brooks, C., *Introductory econometrics for finance. 3rd ed*, Cambridge university press, 2015, P.419- 424.

[6] Bollerslev, T., Marrone, J., Xu, L., Zhou, H., *Stock return predictability and variance risk premia: Statistical inference and international evidence*, Journal of Financial and Quantitative Analysis, 2014, **49**(03), P.633–661. Doi: 10.1017/S0022109014000453.

[7] Campbell, S. D., *Macroeconomic volatility, predictability, and uncertainty in the great moderation*, Journal of Business and Economic Statistics, 2012, **25**(2), P.191-200. Doi: 10.1198/073500106000000558

[8] Campbell, J., Thompson, S., *Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies*, 2008, **21**(4), P.1509–1531. Doi: 10.1093/rfs/hhm055.

[9] Chong, E., Han, C., Park, F., *Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies*, Expert Systems with Applications, 2017, **83**, P.187–205. Doi: 10.1016/j.eswa.2017.04.030.

[10] Diebold, F., Mariano, R., *Comparing predictive accuracy*, Journal of Business and Economic Statistics, 1995, **13**, P.253-263. Doi: 10.1198/073500102753410444.

[11] Ehteshami; S., Hamidian, M., Hajiha, Z., Shokrollahi, S., *Forecasting Stock Trend by Data Mining Algorithm*, 2018, **3**(1), P.97-105. Doi: 10.22034/AMFA.2018.539138.

[12] Enke, D. - Mehdiyev, N., *Stock market prediction using a combination of stepwise regression analysis, differential evolution-based fuzzy clustering, and a fuzzy inference neural network*, Intelligent Automation and Soft Computing, 2013, **19** (4), P. 636–648. Doi: 10.1080/10798587.2013.839287

[13] Farshadfar, Z., Khalili, M., *Study the Effect of Stock Liquidity on Execs Return with Five Factor Arbitrage Pricing Model*, 2016, **9**(29), P.97-109. (In Persian)

[14] Farshadfar, Z., Prokopczuk, M., *Nonlinear Model Improves Stock Return out of Sample Forecasting (Case Study: United State Stock Market)*, Journal of Finance and Managerial Accounting, 2018, **3** (1), P. 1-13.

[15] Ferreira, M., Santa-Clara, P., *Forecasting stock market returns: The sum of the parts is more than the whole. Journal of Financial Economics*, 2011, **100** (3), P.514–537. Doi: 10.1016/j.jfineco.2011.02.003.

[16] Goyal, A., Welch, I., *Predicting the equity premium with dividend ratios*, Management Science, 2003, **49**, P.639-654.

[17] Hammoudeh, S., R., Ewing, B. T., *Relationships among strategic commodities and with financial variables: a new look*, Econ Policy, 2008, **27**(2), P.251–264. Doi: 10.1111/j.1465-7287.2008.00126.x.

[18] Jasemi, M., Kimiagari, A., Memariani, A., *A conceptual model for portfolio management sensitive to mass psychology of market*, International Journal of Industrial Engineering Theory Application and Practice, 2011, **18** (1), P.1–15.

[19] Kim, Y., Enke, D., *Developing a rule change trading system for the futures market using rough set analysis*, Expert Systems with Applications, 2016a, **59**, P.165–173. Doi: 10.1016/j.eswa.2016.04.031

[20] Kim, Y., Enke, D., *Using neural networks to forecast volatility for an asset allocation strategy based on the target volatility*, Procedia Computer Science, 2016b, **95**, P.281–286. Doi: 10.1016/j.procs.2016.09.335

[21] Lee, C., Sehwan, Y., Jongdae, J., *Neural Network model versus ARIMA model in forecasting Korean stock price index*, Information System, 2007, **8** (2), P.372–378.

[22] Lee, K., Chi, A., Yoo, S., Jin, J., *Forecasting Korean stock price index using back propagation neural network model, bayesian chiao's model, and ARIMA model*, Academy of Information and Management Sciences

Journal, 2008, **11** (2), P.53 -75.

[23] Nair, V., Hinton, G., *Rectified linear units improve restricted Boltzmann machines*, 27th international conference on machine learning, 2010, P.807–814.

[24] Narayan, P., Narayan, S., *Modelling the impact of oil prices on Vietnam's stock prices. Applied Energy,* 2010, **87**, P.356-361. Doi: 10.1016/j.apenergy.2009.05.037.

[25] Narayan, P., Sharma, S., *New evidence on oil price and stock returns*, Journal of Banking and Finance, 2011, **35**, 3253-3262. Doi: 10.1016/j.jbankfin.2011.05.010.

[26] Nasr, N., Farhadi Sartangi, M., Madahi, Z., *A Fuzzy Random Walk Technique to Forecasting Volatility of Iran Stock Exchange Index*, 2019, **4**(1), P.15-30. Doi:10.22034/AMFA.2019.583911.1172.

[27] O'Connor, F. A. – Lucey, B. M. – Batten, J. A. – Baur, D. G., *The financial economics of gold a survey. International Review of Financial Analysis*, 2015, **41**, P.186–205.

[28] Phan, D., Sharma, S., Narayan, P., Stock return forecasting: *Some new evidence. International Review of Financial Analysis,* 2015, **40**, P.38–51. Doi: 10.1016/j.irfa.2015.07.005.

[29] Rapach, D., Strauss, J., Zhou, G., *Out-of-sample equity premium prediction: Combination forecast and links to the real economy. Review of Financial Studies,* 2010, **23**, P.821-862. Doi: 10.1093/rfs/hhp063.

[30] Rather, A., Sastry, V., Agarwal, A., *Stock market prediction and Portfolio selection models: a survey. Opsearch,* 2017, **54**, P.558-579.

[31] Soytas, U., Sari, R., Hammoudeh, S., Hacihasanoglu, E., *World oil prices, precious metal prices and macro economy in turkey*, Energy Policy, 2009, **37**(12), P. 5557–5566. Doi: 10.1016/j.enpol.2009.08.020.

[32] Welch, I., Goyal, A., *A comprehensive look at the empirical performance of equity premium prediction,* The Review of Financial Studies, 2008, **21** (4), P.1455-1508. Doi: 10.1093/rfs/hhm014.

[33] Westerlund, J., Narayan, P., *Does the choice of estimator matter when forecasting stock returns*, Journal of Banking and Finance, 2012, **36**, P.2632-2640. Doi :10.1016/j.jbankfin.2012.06.005.

[34] West, K., Asymptotic inference about predictive ability, 1996, Econometrica, **64**, P. 1067-1084.