International Journal of Informatics Media and Communication Technology
ISSN: Online 2682-2881 & Print 2682-2105

Machine Learning Techniques for Predicting the Egyptian Stock Exchange

Heba M. Ezzat
Faculty of Economics and Political Science, Cairo University, Cairo, Egypt
Correspondence should be addressed to heba.ezzat@feps.edu.eg, h.m.ezzat@gmail.com , ***ORCID ID:*** *0000-0002-0786-8000*

DOI: 10.21608/ijimct.2025.230509.1050

Abstract:

Predicting future movement of stock prices is very significant for researchers, investors, and companies. Due to the sensitivity of financial markets to any news, the market absorbs updates instantly. This hardens the forecasting process pretty much. However, efforts exerted in this field reveal that machine learning methods are very promising in predicting stock time series data. Many advances are achieved by implementing different machine learning techniques. Machine learning methods have improved ability of learning and training from datasets to produce more accurate forecasting. In this paper, three different machine learning methods are implemented; Prophet, K-Nearest Neighbor, and Feedforward Neural Network. EGX30, the main index of the Egyptian Stock Exchange, was collected for about 25 years to be modelled by the proposed three methods. To the best of our knowledge, no research has been conducted to predict the Egyptian Stock Exchange with these three models. Accuracy metric was reported to compare the performance of the three models. Surprisingly, Prophet model performed the best though it is not so famous in predicting stock prices.

Keywords

Prophet, K-Nearest Neighbor, Feedforward Neural Network, EGX30, financial time series, deep learning

1. Introduction

Predicting financial markets is very interesting for both researchers and investors. Fama suggests the so-called Efficient Market Hypothesis (EMH) which assumes that financial markets cannot be predicted This is because the prices absorb all the news and this is reflected in financial assets' prices. Many efforts were exerted to predict future prices of financial markets using machine learning techniques [c.f. (Ezzat, 2021), (Kissell, 2020), (Shah, Isah, & Zulkernine, 2019), (Edwards, Magee, & Bassetti, 2018) (Zhong & Enke, 2017) , and (Vaisla & Bhatt, 2010)].

The complex nature of financial time-series could be captured by computational modeling. These complex characteristics are known as stylized facts such as excess volatility, volatility clustering, heavy tailed-distribution of returns, power-law tails, random-walk prices, and fractal structure (Cont, 2001). Computing enhancements enable predicting financial time-series.

There are enormous machine learning techniques that could be exploited to predict stock market prices. Prophet or 'Facebook Prophet' is a machine learning model developed by Facebook to predict univariate time series (Taylor & Letham, 2018). K-Nearest Neighbor (KNN-Regression) is a machine learning algorithm that could be used for classification and finding the regression (Cover & Hart, 1967; Fix & Hodges, 1951). The data set is split to training and test. The model is applied to predict the classification of each new data point to which group. This is found by searching for the nearest neighbor (class). Also, Feedforward Neural Network (FNN) is one of the important deep learning algorithms (Amari, 1972; Hopfield, 1982;

Siegelmann & Sontag, 1991). In this model, the information flows between three layers; input, hidden, and output. The hidden layer is located between the input and output layers. Information is directed to one direction contrary to recurrent neural network, where information could be circulated between the layers. In this research, Prophet, KNN-Regression, and the FNN will be applied.

The three models will be executed on the main index of the Egyptian Stock Exchange; EGX30. As far as we know, no study was run to predict the EGX30 following these three models. The performance of these models will be compared according to their accuracy.

The rest of the paper will include four sections. In Section 2, data collection and pre-processing are described. Section 3 mainly dedicated to explain the prediction models. Major results are illustrated and discussed in Section 4. Finally, Section 5 concludes the paper.

2. Data description and pre-processing

The three models will be applied on the closing price of the main index EGX30 collected from investing.com. The closing prices cover the period from 5/18/1998 to 5/17/2023 with a total number of 6116 observations. Closing prices are used for prediction.

The frameworks followed to implement the three models are displayed in Figure 1.

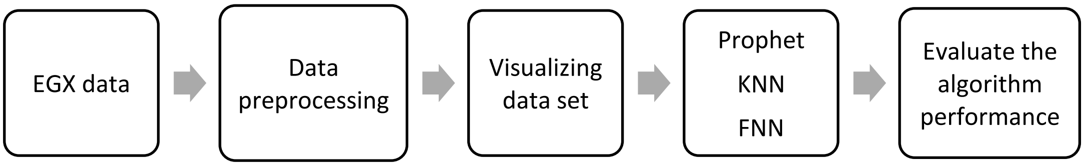

Figure 1. Framework for the Prophet, KNN, and FNN models.

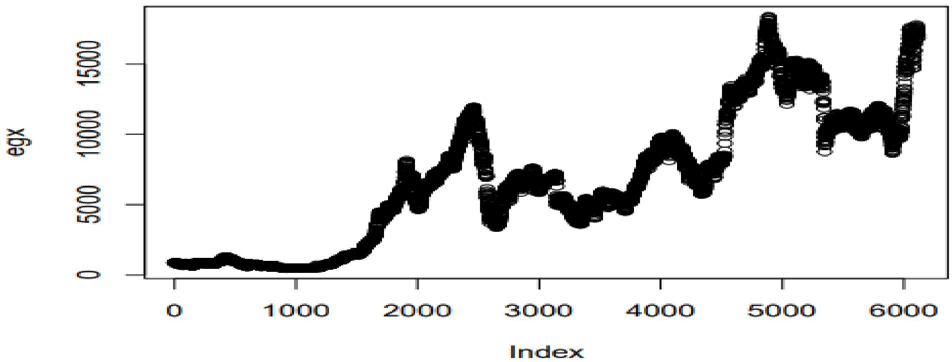

Figure 2. EGX30 closing prices' evolution.

The three suggested models; Prophet, KNN and FNN are following the supervised learning approach. Accordingly, 80% of the data were used for training and 20% for testing and validation. RStudio software was used for implementation. Figure 2 shows the evolution of the EGX30 closing prices over the past 25 years.

3. Prediction models

This research suggests three models; Prophet, KNN, and FNN, for stock price prediction. Before implementing the three models and illustrate their results, a clear description for each model is provided in the following subsections.

3.1 Prophet model

Prophet model is a univariate time-series forecasting model created by Facebook Core Data Science team (Taylor &

Letham, 2018). Prophet works very well in analyzing vast amount of historical data. Prophet is based on splitting the data to additive components, such as trend and seasonality. Prophet implements an additive regression with the form displayed in Eq (1):

$$y_t = g_t + s_t + h_t + \epsilon_t \qquad (1)$$

Where g_t refers to the trend function that models non-periodic changes in the value of the time series, s_t is the periodic seasonality (weekly or yearly), h_t refers to the holiday effect if such information is provided during the modeling stage, finally, ϵ_t represents unobserved factors affecting the dependent variable. A yearly seasonal component is modelled using Fourier series.

There are many advantages of Prophet modelling, such as generating accurate forecasting. Also, Prophet is very fast as it can generate forecasting in seconds. Moreover, Prophet is automatic so that there is no need to perform data preprocessing as it can work with missing data and outliers.

3.2 K-NN regression time-series forecasting

K-NN is one of the most popular machine learning algorithms that are used for classification and modeling regression (Cover & Hart, 1967; Fix & Hodges, 1951). K-NN is a supervised learning method implemented to model non-parametric regression. K-NN regression computes the average of the numerical target of the k-nearest neighbor. K defines the number of neighbors that will be included in computing the average. The nearest neighbors are determined by the Euclidean distance, which is the square root of the sum of squared differences between the corresponding elements of two vectors as illustrated in Eq (2). The smaller the difference, the more similar records will be.

$$Eucledian\ distance = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \qquad (2)$$

Many authors conduct research to predict stock market prices following the K-NN algorithm, such as (Latha, et al., 2022; Alkhatib, Najadat, Hmeidi, & Shatnawi, 2013). The results show that k-NN model performs very well in predicting future stock prices.

3.3 Feedforward Neural Network

Artificial Neural Network (ANN) is a machine learning algorithm aims mainly to mimic the biological nervous system (Amari, 1972; Hopfield, 1982; Siegelmann & Sontag, 1991). Many advances in artificial intelligence were created exploiting ANNs, such as optical recognition, voice recognition, and pattern recognition. There are many types of ANNs. FNN is a unidirectional neural network, where information flows from input, hidden, to output layer.

To understand how FNNs work, consider its simplest form of a single layer perceptron. In this model, a sequence of inputs enter the layer and are multiplied by a set of weights. A sum of the weighted input values is obtained by adding these multiplications together. A threshold is set, usually its value is zero. If the additive value is greater than the threshold, then the value generated is +1. Otherwise, the value generated is -1. This single-layered perceptron model is usually used for classification.

Learning and adaptation are very important aspects of machine learning. In ANN, the delta rule is used, where the output of ANN's nodes are compared to the intended values. The network is adjusting its weights through training to generate more accurate output values. In a multilayered perceptron, the learning process to update weights is almost the same and it is known as backpropagation. In this case, each hidden perceptron is modifying its weights according to the output values generated by the final layer.

Many authors implemented FNNs to predict future stock prices' movements, such as (Jiang, 2021; Namdari & Durrani, 2021; Lee, Siddiqui, Abbas, & Lee, 2017). The results show that FNN is performing very well for stock prediction.

The following section displays the main results we obtained after implementing the three models to the EGX30.

4. Results and discussion

The following illustrates the output obtained from implementing the three models. Figure 3 shows the training and forecasting of EGX30 following the Prophet model. Figure 4 illustrates the separation of the Prophet model to trend and yearly seasonality. Figure 5 displays the training and forecasting of EGX30 following the K-NN model. Finally, Figure 6 shows the training and forecasting of EGX30 following the FNN.

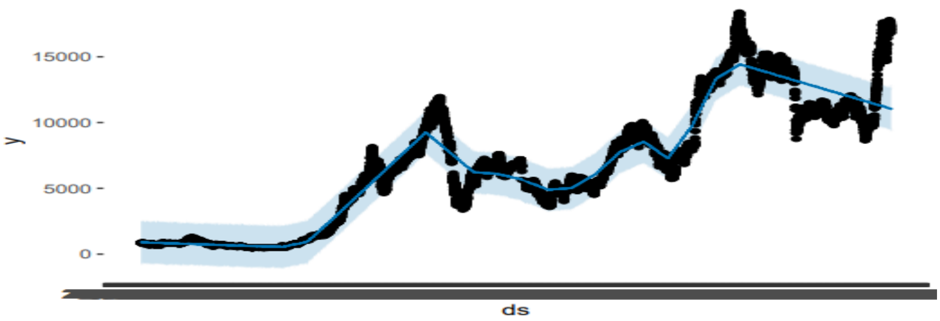

Figure 3. Training and forecasting of data following Prophet model.

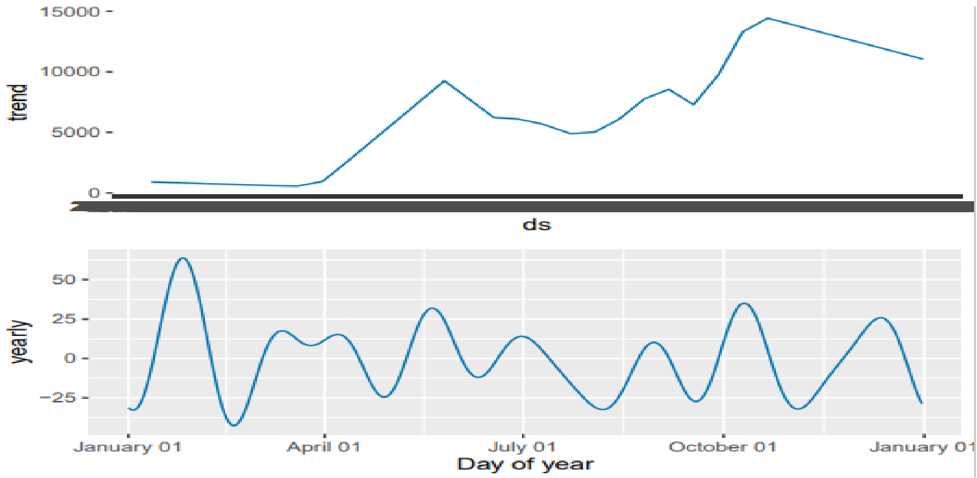

Figure 4. Additive components of Prophet.

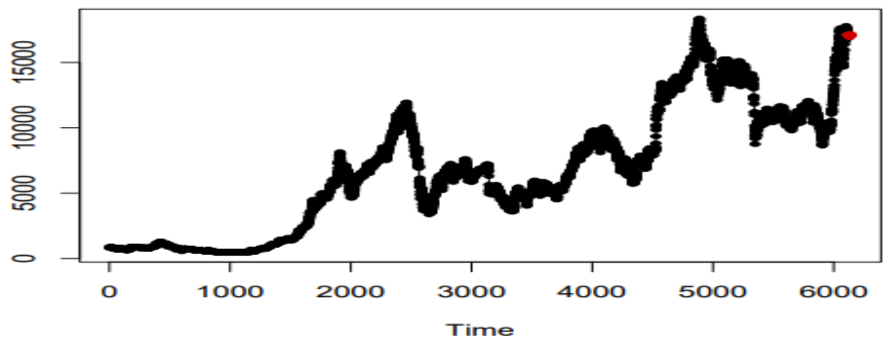

Figure 5. Training and forecasting of data following K-NN model.

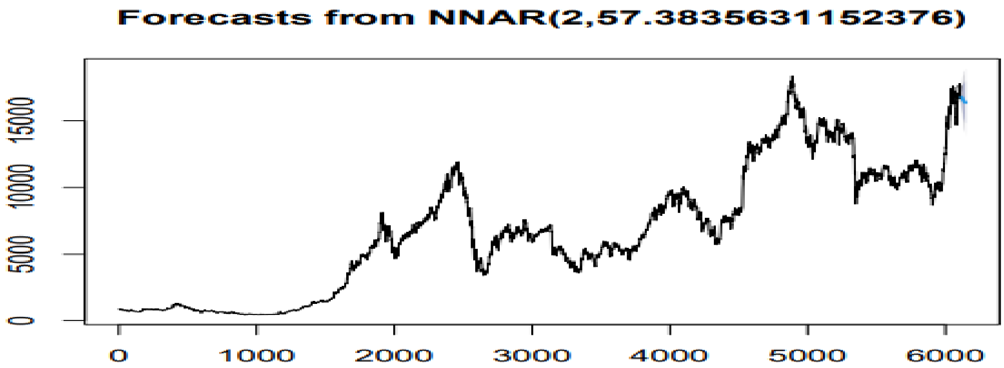

Figure 6. Training and forecasting of data following FNN model.

To compare the performance of the three models, we computed the accuracy metric with three important indicators; Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

RMSE uses the square root of the average square difference between actual values and the predicted ones. MAE defines the average of the absolute deviation of actual values from their respective predictive values. MAPE is the average of absolute percentage errors of each entry in a dataset to compute the accuracy of forecasting comparing to the actual values. To compute these three accuracy measures, we followed Eq. 3 to Eq. 5. The results for the three models are summarized in Table 1.

$$RMSE = \sqrt{\left(\frac{1}{n}\right) * \sum_{t=1}^{n} (y_t - \hat{y}_t)^2} \tag{3}$$

$$MAE = \left(\frac{1}{n}\right) * \sum_{t=1}^{n} |y_t - \hat{y}_t| \tag{4}$$

$$MAPE = \left(\frac{1}{n}\right) * \sum_{t=1}^{n} \left(\frac{|y_t - \hat{y}_t|}{|y_t|}\right) * 100 \tag{5}$$

Table 1. Accuracy measures for the performance of the three models.

Model	RMSE	MAE	MAPE
Prophet	22.2108	18.2503	0.7939
K-NN	1125.0607	950.0183	5.4923
FNN	110.9061	69.1803	1.1183

Notice that, Prophet performed the best while K-NN is considered the worst.

5. Conclusion

Forecasting future behavior of stock markets is very critical for investors and companies. Due to the complexity of financial time series, forecasting is considered very difficult. Computational modelling enabled forecasting future movements of stock markets. Different machine learning techniques were exploited over the last decade to learn and train from the dataset. In this paper, three different machine learning techniques were implemented; Prophet, K-NN, and FNN. The accuracy metric for the three models is reported. Unexpectedly, Prophet provides the least RMSE, MAE, and MAPE. Thereby, Prophet is considered to be the best fit model for these data.

References

Alkhatib, K., Najadat, H., Hmeidi, I., & Shatnawi, M. K. (2013). Stock Price Prediction Using K-Nearest Neighbor (kNN) Algorithm. *International Journal of Business, Humanities and Technology, 3*(3), 32 - 44.

Amari, S. ..-I. (1972). Learning Patterns and Pattern Sequences by Self-Organizing Nets of Threshold Elements. *S. . -I. Amari, "Learning Patterns and Pattern Sequences by Self-Organizing Nets of IEEE Transactions on Computers, C-21*(11), 1197-1206. doi:10.1109/T-C.1972.223477

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance, 1*(2), 223 - 236.

Cover, T., & Hart, P. (1967). Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory, 13*(1), 21–27.

Edwards, R., Magee, J., & Bassetti, W. (2018). *Technical Analysis of Stock Trends.* Boca Raton: CRC Press.

Ezzat, H. (2021). Principal Component Regression for Egyptian Stock Market Prediction. *The International Journal of Informatics, Media and Communication Technology, 3*(1), 23 - 43. doi:10.21608/ijimct.2021.169612

Fix, E., & Hodges, J. (1951). *Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties.* Texas: USAF School of Aviation Medicine, Randolph Field.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, 79*(8), 2554 - 2558.

Jiang, W. (2021). Applications of deep learning in stock market prediction: Recent progress. *Expert Systems with Applications, 184*, 1 - 21. doi:10.1016/j.eswa.2021.115537

Kissell, R. (2020). *Algorithmic Trading Methods: Applications Using Advanced Statistics, Optimization, and Machine Learning Techniques.* Academic Press.

Latha, R. S., Sreekanth, G. R., Suganthe, R. C., Geetha, M., Selvaraj, R. E., Balaji, S., . . . Ponnusamy, P. P. (2022). Stock Movement Prediction using KNN Machine Learning Algorithm. *2022 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1 - 5). Coimbatore, India: IEEE Xplore. doi:10.1109/ICCCI54379.2022.9740781

Lee, X., Siddiqui, M., Abbas, A., & Lee, S. U.-J. (2017). A Feed-forward Neural Network Model for Predicting Stock Price. *Proceedings of the Korean Information Science Society Conference*, (pp. 1018 - 1020). South Korea.

Namdari, A., & Durrani, T. S. (2021). A Multilayer Feedforward Perceptron Model in Neural Networks for Predicting Stock Market Short-term Trends. *SN Oper. Res. Forum, 38*(2), 1 - 30. doi:10.1007/s43069-021-00071-2

Shah, D., Isah, H., & Zulkernine, F. (2019). Stock market analysis: A review and taxonomy of prediction techniques. *Int. J. Financ. Stud., 7*(2), 1 - 22.

Siegelmann, H., & Sontag, E. (1991). Turing computability with neural nets. *Appl. Math. Lett., 4*(6), 77 - 80. doi:10.1016/0893-9659(91)90080-F

Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *The American Statistician, 72*(1), 37-45. doi:10.1080/00031305.2017.1380080

Vaisla, K., & Bhatt, A. (2010). An analysis of the performance of artificial neural network technique for stock market forecasting. *Int. J. Comput. Sci. Eng., 2*(6), 2104 - 2109.

Zhong, X., & Enke, D. (2017). Forecasting daily stock market return using dimensionality reduction. *Expert Syst. Appl., 67*, 126 - 139.