PRIMARY RESEARCH

Prediction of Stock Price of GCC Islamic Banks Using Neural Network

Sakina 1*, Muhamamd Anees Khan 2
1, 2 Bahria University, Islamabad, Pakistan

Keywords
Artificial Intelligence
Artificial Neural Network
K-Nearest Neighbor
Stock Prediction
MAPE
RMSE
MBE

Abstract.

Purpose: This study aims to evaluate the predicting ability of two machine learning models (ANN and KNN) to predict the stock prices of GCC Islamic banks.

Methodology: The study analyzes the accuracy rates of two machine learning models to predict the stock prices of GCC Islamic banks by using the daily stock prices (open, close, high, and low) for a period of ten years (May 4th, 2012–May 31st, 2022).

Findings: The predicting ability of the trained ANN and KNN models is evaluated using standard strategic indicators: root mean square error (RMSE), mean bias error (MBE), and accuracy rates.

Originality: This study is distinctive in that it emphasizes the significance of machine learning models as an alternative to traditional methods for predicting stock prices.

Research Limitations: The study has conducted a comparative analysis of the accuracy rates of only two machine learning models and determined the superior model among them.

Practical Implication: The findings of this study have certain implications, as the results can be used to gain significant profits in the GCC region, as they help in predicting market fluctuations, discovering hidden relationships in the data sets, and making accurate investment decisions.

KAUJIE Classification: L41, L42, L43
JEL Classification: G14, G15, G17

*Corresponding author: Sakina
†Email: sakina.ali835@hotmail.com; ORCID: https://orcid.org/0009-0007-1802-9687

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License http://creativecommons.org/licenses/by/4.0/
INTRODUCTION

Islamic financial markets are expected to play a prominent role in the Shariah compliant stock market by providing investment opportunities through their financial products (Forte & Miglietta, 2007). After the global financial crisis in 2008, the financial and investment markets went through a significant period of change. However, the Islamic financial markets have made efforts to demonstrate their ability to sustain their performance in this new environment (Omar Farooq & Hasib Reza, 2014). The recent development of Islamic financial instruments and the establishment of Shariah compliance have led to a significant rise in investment and investor confidence in these financial instruments (Mnif et al., 2020). As a consequence, return on Shariah-compliant assets has steadily increased (Al-Rifai, 2012). Additionally, compared to conventional performance metrics, Islamic indicators are more transparent, which has resulted in the distribution of good returns (Ousama et al., 2020).

Likewise, the recent influx of foreign investment into Islamic financial markets can be attributed to the rapid development of Islamic nations, specifically those comprising the Gulf Cooperation Council (GCC), which is primarily driven by the accumulation of hydrocarbon wealth (Platonova et al., 2018). The literature states that in order to achieve long-term growth utilizing Islamic financial products, nations must provide investment alternatives that align with advancements observed in significant international financial centres. Financial intermediaries have been motivated to offer services and products as a result of the Islamic capital markets’ ongoing expansion and exceptional growth. According to Ousama et al. (2020), Islamic financial institutions are characterized by a sense of social and moral obligation to manage investments. Islamic finance has gained liquidity in regional capital markets by providing investment options that adhere to Islamic law, in parallel with advancements in the Islamic capital market. Over the past few years, the marketplaces in the GCC countries have experienced consistent growth, outperforming the global average. Likewise, the present decrease in oil prices is compelling these nations to expand and strengthen their markets (Mezghani & Boujellbène, 2018).

This increasing growth in muslim countries has attracted investors to emphasize the significance of acquiring a comprehensive understanding of these markets and making investments. The investors in these markets need an accurate understanding of future pricing to secure their investments. Accurately predicting fluctuations in stock prices can optimize investors’ returns and motivate them to increase their investments (Abedifar et al., 2020). Traditional financial theories argue that predicting stock prices with any degree of accuracy is impossible, there are formal reasons to the contrary. They demonstrate that, given the correct variables and modelling approaches, future stock prices and stock price movement patterns may be predicted rather accurately (Mehtab et al., 2021). The existing traditional methods of prediction (fundamental, technical and time series) lack the ability to account for the dynamics and interdependencies among the stocks due to the complexities, uncertainty, and non-linearity of the time series data, which has shifted the focus to non-linear models (Selvin et al., 2017; Vijh et al., 2020). The most powerful non-linear models are machine learning models (Support vector Machine, Artificial Neural Network, K-Nearest Neighbor, Multilayer perception, Neural Networks, Naive Bayes, Long Short-term memory, and Convolutional...
Neural Networks) which are used to process sequential data and help find hidden patterns and underlying dynamics that can be used to predict stock price. These models have an accuracy rate of 60–86% prediction earlier (Anand, 2021; Khan et al., 2022; Mehtab et al., 2021; Selvin et al., 2017). Considering the importance of investments in GCC regions and the need to accurately predicting the stock prices, this study mainly focuses on comparing the performance of two machine learning (ML) models (ANN and KNN) to predict the future values of the stock prices of GCC Islamic banks in order to help investors predict and earn promising returns. It helps the investors identify which particular model can best be used to predict stock prices in the GCC region. This can reduce stock return uncertainty, which enables investors or institutions to make investment decisions more wisely and achieve promising outcomes in this region.

The results of the study are beneficial to foreign investors, bankers, brokers, and everyone else who is interested in purchasing Islamic equities and forecasting the value of companies in the Gulf Cooperation Council. By employing artificial neural network (ANN) and k-nearest neighbors (KNN) techniques, individuals can utilize it to determine the optimal moment for investment and maximize their earnings. In order to limit the risk of an excessively volatile stock market and create smooth policies that produce significant profits for all investors, policymakers may also use the application of these models. Finally, it adds value by highlighting the importance of ML models, which can enable researchers to concentrate on more advanced prediction approaches rather than more conventional ones. It will advance knowledge of non-linear financial time series in academia.

**LITERATURE REVIEW**

In his efficient market hypothesis, Markowitz posited that financial markets necessitate optimal allocation, extensive diversification, and transparency. Nevertheless, the literature on Islamic financial markets is still being developed because of their relatively short operational history. The primary objective of Islamic stock markets is to facilitate the transfer of funds from entities with financial surpluses to entities with deficits while adhering to financial transactions that are free from riba. Deficit units refer to business-sector enterprises that seek to raise equity money in order to finance new funding. Nevertheless, Islamic stock capital financing and investments have some operational and financial risks that require attention. For decades, academics have been interested in the predictability of these stock prices. Financial analysts, brokers, and scholars all have faith in the predictability of stock market movements based on statistical methodology and historical data. Several academic studies have been conducted to identify historical trends in financial time series. A massive quantity of historical data is now available for examination due to the advancement of storage and tracking technologies during the last 20 years. As a result, machine learning techniques have become the primary focus of new works (Lachiheb & Gouider, 2018). There are papers suggesting that stock prices can be predicted using multiple machine learning models, but there is no precise consensus on which specific model better predicts the prices of different
regions. Some of the machine learning models used to predict stock prices are Artificial Neural Network (ANN) and K-nearest Neighbor (KNN) (Dash & Dash, 2016).

The popularity of ANN as a predicting model originates from their adeptness at recognizing and interpreting relationships among nonlinear variables. Moreover, ANNs excel in handling extensive datasets with rapid and frequent fluctuations, making them an optimal choice for predicting stock market trends. Numerous studies have aimed to enhance the precision and computational efficiency of stock price (SP) prediction, with ANNs standing out as sophisticated algorithms tailored to address complex challenges beyond the scope of traditional neural networks or basic machine learning models (Wu et al., 2001). These networks boast a more intricate and interconnected structure compared to the human brain, utilizing algebraic equations as a foundational framework to steer data toward modelling time-series or specific patterns (Dhenuvakonda et al., 2020). Likewise, Vijh et al. (2020) evaluated the accuracy of Artificial Neural Networks (ANNs) against Random Forest by employing evaluation metrics like MAPE, RMSE, and MBE, validating the superior performance of ANNs over Random Forest.

To enhance stock price prediction, some authors propose employing a hybrid strategy. Adebiyi et al. (2012) integrated technical and fundamental aspects of social media indicators with publicly available stock data, resulting in significantly improved outcomes compared to relying solely on technical analysis. Traders and investors found the hybridized forecast reliable for making qualitative decisions. In a separate study, Moghaddam et al. (2016) explored an artificial neural network’s (ANN) capability to forecast the daily NASDAQ stock exchange values. They evaluated several feed-forward ANNs trained via back propagation, creating a comprehensive model using daily NASDAQ stock prices. Splitting the data into training (first 70 days) and testing (remaining 29 days), they assessed predictive ability using determination coefficient ($R^2$) and mean square error (MSE) metrics. Interestingly, the model’s results showed no noticeable variation in performance based on the four or nine previous working days as input parameters.

In their study, Vijh et al. (2020) utilized existing variables (low, high, open, and close) to generate new variables (high-low price, close-open, 7-day moving average, 14-day moving average, 21-day moving average, and 7-day standard deviation) aimed at improving the accuracy of predicted values. Through assessments using model evaluation tools (RMSE, MAPE, and MBE), their study clearly demonstrated the superiority of ANN over Random Forest in predicting SP. Similarly, Shahvaroughi Farahani & Razavi Hajiagha (2021) employed advanced metaheuristic methods like the bat algorithm (BA) and social spider optimization (SSO) to train artificial neural networks (ANN) using various technical indicators as input variables. Comparing the results using time series models, specifically ARMA and ARIMA models, both indicated that the ANN model exhibited better predictability than the time series models based on the MAE outcomes.

The stock market offers an ideal setting for data mining and business exploration, given its extensive and constantly evolving information sources. Researchers utilized the k-nearest neighbor method in conjunction with a non-linear regression technique to predict stock values for six prominent companies listed on the Jordanian stock exchange, aiming to support
informed decisions by investors, managers, decision-makers, and users. Their findings highlight the reliability of the KNN algorithm, showcasing low error rates and consistently logical and satisfactory prediction outcomes (Alkhatib et al., 2013). Additionally, machine learning algorithms like KNN, SVM, neural networks, Bayesian networks, and AdaBoosts were applied to estimate stock transaction volumes for ten listed firms on the Karachi stock exchange, demonstrating encouraging results with minimal error, as indicated by the lowest MAE of 0.0904 and RMSE of 0.2402 (Ghazanfar et al., 2017).

A fusion model, blending Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) with various kernel functions, was applied to forecasting profit or loss. The SVM output aids in identifying the most precise nearest neighbor from the training set subsequently used to forecast the future stock values of indices like the Bombay Stock Exchange (BSE Sensex) and CNX Nifty. Evaluating this model involves metrics such as Root Mean Square Error (RMSE), Mean Squared Forecast Error (MSFE), and Mean Absolute Forecast Error (MAFE). Nayak et al. (2015) highlighted that this SVM-KNN combined model significantly enhances stock analysis by amalgamating SVM training and KNN analysis. Similarly, in accurately predicting stock market indices, Chen & Hao (2017) also propose a straightforward hybrid framework, utilizing feature-weighted SVM and feature-weighted KNN. The value of each feature is determined through an evaluation of information gain employing weighted theory. Using feature-weighted KNN, they predict future stock market indices by approximating the k-weighted closest neighbors from historical datasets. Their study, assessing Shanghai and Shenzhen stock exchange indices, confirms the method’s accuracy in predicting indices across short, medium, and long-term periods.

Several studies have compared various machine learning models to gauge their effectiveness in predicting stock market trends. These models encompass a range of methodologies including SVM, Random Forest, KNN, Naive Bayes, and Soft Max. The empirical findings illustrate that the Naive Bayesian Classifier shows superior performance in small datasets, whereas the RF technique outperforms it in larger datasets. Notably, the precision of each method tends to decrease as the number of technical indicators increases (Kumar et al., 2018). In a similar way, Lawal et al. (2020) conducted a comparative analysis, finding Support Vector Machines (SVM) as the frequently used method for stock price prediction due to their remarkable performance and accuracy. Additionally, approaches like ANN, KNN, Naive Bayes, Random Forest, Linear Regression, and SVR showed promising prediction outcomes. Lately, Madeeh & Abdullah (2021) conducted a comparison between the KNN and Random Forest algorithms to determine the superior predictive model. Their study, involving data preprocessing, model implementation, and efficacy evaluation, indicated high accuracy in both models, with the RF model exhibiting the highest prediction accuracy (93.23%, 93.12%, and 93.17%, respectively). The evaluation was based on precision, recall, and the F-measure metrics.
RESEARCH METHODOLOGY

The daily historical prices (Henrique et al., 2018; Kumar et al., 2016) of 34 Islamic banks for a period from May 4th, 2012, to May 31st, 2022, are used to predict the stock price movement in this study. The selected index is based on the availability of data and its impact on the economy. The data feature contains information about stock prices, such as high, low, open, and close. The study uses Python 3.10 for splitting the data into training and testing, applying the machine learning model, and evaluating the model to attain its main objective. Figure No. 1 explains the data analysis technique used in this study.

![Machine learning Process](image)

**FIGURE 1.** Flow Chart of ML Model

The raw data is imported into the programming language Python to assess the quality and develop a model to predict the stock prices of the world’s top indices. In order to do so, the Python language needs some libraries to be imported. The libraries are NumPy, Pandas, Matplotlib, Seaborn, Sklearn, multi-layer regressors, and metrics.

### Data Preprocessing

The data is initially imported into Python to be cleaned by identifying missing and duplicate values. Data cleaning is the act of eliminating inherent mistakes that may bias our study data and make it less effective. Cleaning can be done in a variety of ways, including finding missing values, deleting empty cells or rows, removing duplicate values, eliminating outliers, and normalizing inputs. Its goal is to guarantee that there are no, or as few as feasible, flaws that might impact our final analysis and accuracy level. All stock price data has been imported into Python, and no missing or duplicate values were discovered while cleaning and filtering the study’s data.
The above Table No. 1 explains the descriptive statistics of the data used in the study. It clearly states the data features (open, low, high, and closing prices) of GCC Islamic banks. The high values of the standard deviation of all the features used in the study confirm the fluctuation in the prices, indicating the nonnormality of the data used. This volatility in prices during the study period can be seen in the graph below.

**FIGURE 2.** Historical Prices

### Data Splitting
The preprocessed and scaled data is divided into train and test groups using the train-test split function. The train test model is used to validate the process of stimulating our model to work with new data. Data mining techniques imply the division of the data set into training and testing. The training data is used to train the ML model and predict according to the algorithm trained, while the testing data is used to evaluate the ML model’s prediction accuracy. The size of the train and test sets has been explicitly divided into 70% and 30% in this study, where 70% of the data is trained for the model training and the remaining 30% is used to check the predicting ability of the trained model.
Machine Learning Models
In this study, the ML algorithms employed were ANN and KNN. The following subsections provide a brief description of each algorithm.

Artificial Neural Network
ANN is a biologically inspired subfield of artificial intelligence based on the structure of the human brain. ANN neurons, like human neurons, are interconnected at numerous levels. These neurons are known as nodes. It aims to duplicate the human brain’s network of neurons, which is used to comprehend information and make judgments in a human-like manner. This method seeks to train computers to behave similarly to linked brain cells. It is made up of three major layers. Input, output, and a hidden layer. Artificial input neurons make up a neural network’s first layer, which receives the initial data that later layers will process. It accepts input in a number of different forms. Artificial neurons in the hidden layer receive a set of weighted inputs and create an output using an activation function. The last layer is where the expected results are obtained (Chhajer et al., 2022).

The above figure represents the ANN process, where the inputs \(a_1, a_2, a_3, \ldots, a_n\) are multiplied by their respective weights and passed to other layers through activation functions to get the desired output.

\[
\text{Sum of input} = a_1 \times w_1 + a_2 \times w_2 + a_3 \times w_3 + \ldots + a_n \times w_n + b.
\]

Where “a1”, “a2”, “a3” and “an” are the input values (high, low, and open prices), “w” represents the weights of each input, and “b” is the bias term. The sum of inputs is passed to the final layer with a threshold function called an activation function. It is an ANN function that provides a small value for tiny inputs and a larger value when the inputs exceed a particular threshold. They decide if the neuron should be activated or not. This study uses the Rectified Linear Activation Function (Relu) to get the desired output.

![Artificial Neural Network](image-url)

FIGURE 3. Artificial Neural Network
K-Nearest Neighbor (KNN)

KNN is a supervised, non-parametric, controlled learning approach that employs proximity to categorize or anticipate how a group of individual data points will be arranged. It can tackle both classification and regression issues. The algorithms label the data depending on the nearest neighbors after computing each distance (Yunneng, 2020). Euclidean distance, Manhattan distance, and Minkowski distance are the three most often used distance computation methodologies (Imandoust et al., 2013; Uludag & Gursoy, 2020). This study used the most commonly used technique “Euclidean Distance” to calculate distance based on the following formula.

$$\sqrt{\sum_{i=0}^{n} (x_i - x)2}$$ (2)

The distance between two points is calculated by taking the square root of the sum of their differences.

The anticipated value is dependent on the size of the chosen k values and the distance calculation technique. After determining the size of the neighbor and the distance calculation method, we can make predictions based on the average value of the k nearest neighbors using the formula below.

$$y = \frac{1}{k} \sum_{i=1}^{k} y_i$$ (3)

Yi represents the quantity of case examples, while Y represents the expected outcome of the model. The study employs Euclidean distance, as suggested by Imandoust et al. (2013) to measure the distance between the neighbor values with a k value of 5 (Ledhem, 2022).

ML Model Performance Evaluation

In order to evaluate the model’s performance, multiple assessment criteria are employed. It helps to assess the effectiveness of the model and identify its strengths and limitations. In order to evaluate the performance of the artificial neural network (ANN) and KNN, this study employed the root mean square error (RMSE), mean bias error (MBE), and accuracy metrics.

Root Mean Square (RMSE)

RMSE is a commonly employed metric for evaluating the predictive accuracy of machine learning models. Euclidean distance is employed to quantify the extent of deviation between the expected values and the actual values.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x^{(i)} - y^{(i)})^2}$$ (4)

where N is the number of data points, $x^{(i)}$ is the i-th measurement, and $y^{(i)}$ is its corresponding prediction.

Mean Bias Error

Mean bias error is used to calculate the mean difference between predicted and actual values. As a result, the lower figures are significant (Ledhem, 2022; Vijh et al., 2020).
MBE can be calculated as,

$$\text{MBE} = \frac{1}{n} \Sigma_{i=1}^{n} (x(i) - y(i))$$  \hspace{1cm} (5)

$x(i)$ is the i-th measurement, and $y(i)$ is its corresponding prediction.

**Accuracy**

Accuracy is the state of being right and precise in prediction. It is a critical and extensively used performance evaluation indicator in a machine learning model (Madge & Bhatt, 2015).

**RESULTS AND DISCUSSION**

The table below presents the lowest values of RMSE, MBE, and accuracy rate for the sample data collected for Islamic banks in the GCC region. The table confirms that both models perform well in predicting stock price movements. However, the predicting ability of KNN is highest for the collected data, with the highest predictive accuracy level of 99.96% and lower values of RMSE and MBE. This result has shown comparatively better accuracy compared to previous studies, with 77.6% accuracy in DJIA, 76% in S&P 500, and 74.4% accuracy in NASDAQ (Shahvaroughi Farahani & Razavi Hajiagha, 2021).

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>MBE</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>0.01052</td>
<td>0.00053</td>
<td>99.81%</td>
</tr>
<tr>
<td>KNN</td>
<td>0.00466</td>
<td>0.00003</td>
<td>99.96%</td>
</tr>
</tbody>
</table>

**FIGURE 4.** ANN Epochs vs Error Rate

To achieve the required degree of accuracy with the lowest possible error rate, numerous epochs were used during the training the data throughout the forward and backward propaga-
ition procedures to alter the weights of the data’s input characteristics. Epoch frequency is a hyperparameter that determines how many times the learning algorithm must analyse the whole data set, which is determined by the amount and complexity of the data. The number of epochs may enhance precision up to a point, after which it suffers from data overfitting. As a result, the data is trained over numerous ideal epochs to achieve the lowest error variance, allowing the data to predict with the maximum accuracy.

The graphs represent the relationship between the number of epochs and the error rate that is used to reach the minimum error variance and highest accuracy rate of ANN model for Islamic Banks in GCC region.

**Conclusion**

A country’s financial stability heavily relies on the income generated by today’s stock markets, as businesses rely substantially on the revenues generated by these markets, making market analysis critical. However, predicting stock price changes has proven to be difficult due to the combined use of multiple factors. Consistent swings in stock prices raise the risk of investment and economic losses and eventually impact the country’s economic structure. However, correct stock price prediction may provide investors with numerous possibilities for increasing the value of their assets. Machine learning models are game changers, as previous research has shown that these models can forecast with high accuracy.

This paper basically aimed to apply the neural machine learning models to predict the daily stock prices of GCC Islamic banks using the historical prices of high, low, open, and close for a period from May 4th, 2012, to May 31st, 2022. According to the ML (RMSE, MBE, and Accuracy) regression findings, ANN can predict the closing prices with 99.81% and KNN with 99.96%. Model evaluation statistics, RMSE, MBE, and accuracy demonstrate that KNN can predict this index with relatively low errors. As a result, the study can recommend using the KNN model while making investment decisions in GCC Islamic banks. These findings may aid high-frequency traders in improving stock return and covariance prediction in this market. Future studies may focus on using other machine learning models and comparing their accuracy levels to identify the best ML model to help investors invest in this market.

**References**


