The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1746-5664.htm

Predicting daily precision improvement of Jakarta Islamic Index in Indonesia's Islamic stock market using big data mining

Mohammed Ayoub Ledhem
Department of Economics, University Centre of Maghnia, Maghnia, Algeria, and
Warda Moussaoui
Department of Economics, Dr Yahia Fares University of Medea, Medea, Algeria

Received 19 December 2022
Revised 4 April 2023
22 July 2023
Accepted 3 September 2023

Abstract

Purpose – This paper aims to apply several data mining techniques for predicting the daily precision improvement of Jakarta Islamic Index (JKII) prices based on big data of symmetric volatility in Indonesia's Islamic stock market.

Design/methodology/approach – This research uses big data mining techniques to predict daily precision improvement of JKII prices by applying the AdaBoost, K-nearest neighbor, random forest and artificial neural networks. This research uses big data with symmetric volatility as inputs in the predicting model, whereas the closing prices of JKII were used as the target outputs of daily precision improvement. For choosing the optimal prediction performance according to the criteria of the lowest prediction errors, this research uses four metrics of mean absolute error, mean squared error, root mean squared error and *R*-squared.

Findings – The experimental results determine that the optimal technique for predicting the daily precision improvement of the JKII prices in Indonesia's Islamic stock market is the AdaBoost technique, which generates the optimal predicting performance with the lowest prediction errors, and provides the optimum knowledge from the big data of symmetric volatility in Indonesia's Islamic stock market. In addition, the random forest technique is also considered another robust technique in predicting the daily precision improvement of the JKII prices as it delivers closer values to the optimal performance of the AdaBoost technique.

Practical implications – This research is filling the literature gap of the absence of using big data mining techniques in the prediction process of Islamic stock markets by delivering new operational techniques for predicting the daily stock precision improvement. Also, it helps investors to manage the optimal portfolios and to decrease the risk of trading in global Islamic stock markets based on using big data mining of symmetric volatility.

The authors would like to thank Laboratory of Quantitative Economics Applied to Development "Laboratoire d'Economie Quantitative Appliquée au développement (LEQAD)" and the General Directorate of Scientific Research and Technological Development "La Direction Générale de la Recherche Scientifique et du Développement Technologique (DGRSDT)" under the Algerian Ministry of Higher Education and Scientific Research "le Ministère de l'Enseignement Supérieur et la Recherche Scientifique (Algerie)" for sponsoring this research.

The authors would also like to thank Pr. Dr. Tarek Djeddi, Head of the Department of Quantitative Economics and Foresight in the National Higher School of Statistics and Applied Economics, Algeria.

Since acceptance of this article, the following author has updated his affiliation: Mohammed Ayoub Ledhem is at the Department of Quantitative Economics and Foresight in the National Higher School of Statistics and Applied Economics "Ecole Nationale Supérieure de Statistique et d'Economie Appliquée (ENSSEA)", Kolea, Algeria.

Journal of Modelling in Management
Vol. 19 No. 3, 2024
pp. 765-786

1746-5664
DOI 10.1108/JM2-12-2022-0291

Originality/value – This research is a pioneer in using big data mining of symmetric volatility in the prediction of an Islamic stock market index.

Keywords Data mining, Data analysis, Financial analysis

Paper type Research paper

1. Introduction

Predicting the trends in the stock market index is considered a significant aim in the financial sector as a relatively exact prediction has the ability to increase stock market profits, generate larger financial advantages and defend against market uncertainties (Livieris *et al.*, 2019). The stock market is known for its high amounts of volatility and inconsistency and investors are always seeking a precise and effective strategy to manage their stock investments (Wang *et al.*, 2020). Financial analysis has undoubtedly evolved from a much more qualitative toward a more quantitative discipline, which is also dependent on the harvesting of facts from databases. Artificial intelligence (AI) methods are becoming useful analytical techniques for predicting during recent years by assessing and exploiting information derived from financial data (Jannes, 2018).

According to Aslam *et al.* (2021), the use of data mining techniques and other new technologies like AI, deep learning (DL) and machine learning (ML) in predicting the trends of Islamic stock markets has become critical for investors to make appropriate investment choices, particularly for the Jakarta Islamic Index (JKII) in Indonesia' Islamic stock market, which has experienced high volatility and unsteady movements (Irsalinda *et al.*, 2020; Ledhem, 2022).

Indonesia's Islamic stock market has grown to become one of the most global sophisticated Islamic stock markets in the world, with excellent profit potential (Qizam *et al.*, 2020). The construction of Islamic stock markets in Indonesia has become large, whereas investments in Islamic stock markets had first been launched in 1997; afterward, in 2000, the JKII showed up to inspire investors who would like to increase profits from the investment options in Islamic stock markets (the JKII is the first Islamic stock index, which contained 30 stocks from Indonesia' Islamic stock market) (Subekti *et al.*, 2020).

In the practice of Islamic stock investments, Islamic stock markets are similar to standard ones; the difference is that the procedure and impact of trading do not breach the Islamic law rule prohibiting interest rates (Subekti *et al.*, 2020). The JKII, like other Islamic stock market indices, has a unique set of features that distinguish it from conventional ones in comparison to risk, profit and loss sharing and informational effectiveness. On the one hand, Islamic stock indices are riskier compared to conventional counterparts as a result of a lack of diversity (Albaity and Ahmad, 2008). Because Islamic finance adheres to the profit and loss sharing rule, Al-Zoubi and Maghyereh (2007) established that the Islamic stock index is less volatile than the regular. Furthermore, Ben Rejeb and Arfaoui (2019) exposed that the Islamic stock market index is more informationally effective than the standard one. In terms of volatility, nevertheless, Ben Rejeb and Arfaoui (2019) showed that Islamic stock indexes are much more volatile than their other standard counterparts. Because this study is concentrating on the element of volatility in the JKII index, Irsalinda *et al.* (2020) and Ledhem (2022) show that the JKII index is extremely volatile and has seen unbalanced trends.

As stated by Irsalinda *et al.* (2020) and Ledhem (2022), there is an absence of newer statistical methodology in Indonesia's Islamic stock markets, especially in predicting JKII price volatility, and because the JKII was experiencing volatility and unstable behavior, investors were forced to use modern techniques such as data mining, DL, AI and ML in predicting JKII price movements. Consequently, as an expansion of the research of Irsalinda *et al.* (2020) and Ledhem (2022), which focuses only on using artificial neural networks (ANNs) in the Islamic capital market of Indonesia, this research aims to use other efficient techniques of data mining for predicting daily precision improvement of JKII price movements and to assist investors in achieving their investment objectives and making precise investment decisions in the Islamic stock markets.

The suitable approach for investigating movements of stock market volatility framework is to characterize it from the perspective of its information distribution pattern across a conditional volatility analysis (Ampomah *et al.*, 2021; Vijh *et al.*, 2020). In practice, asset price volatility in the stock market is ultimately driven by data information (Ampomah *et al.*, 2021; Khedr and Yaseen, 2017; Nayak *et al.*, 2015; Vijh *et al.*, 2020). The stock market's instability is symmetric data that can be sustained in the future (Othman *et al.*, 2020). Furthermore, symmetric volatility is thought to be more robust to historical (lagged) values than existing market value shocks (Othman *et al.*, 2020). Thus, in this research, JKII's symmetric volatility framework is used to predict its daily price trend according to the symmetric volatility information (that perfectly demonstrates supply and demand details). This symmetric volatility is determined by four input price attributes:

(1) close price (CP);
(2) low price (LP);
(3) high price (HP); and
(4) open price (OP).

Hence, this research aims to afford robust and accurate approaches, such as data mining techniques to predict the daily precision improvement of JKII prices using the symmetric volatility data as inputs in the prediction model for helping investors to achieve their investment objectives by decreasing the stock risk exchange and to make the managing investment decisions in Indonesia' Islamic stock market. Moreover, one of the modern technologies for improving prediction performance is big data mining (Yang, 2021). Therefore, this research is applying big data mining of symmetric volatility for the prediction process to achieve the optimal prediction performance of daily precision improvement of JKII prices.

Unlike the existing research, this research contribution is the prediction of JKII prices with big data mining of symmetric volatility for improving the daily precision of the JKII prices. Besides, this research investigates the most appropriate big data mining technique for predicting the daily precision of the JKII prices according to the criteria of the optimal performance score of testing and training. Consequently, this research answers the following question, "What is the optimal big data mining technique for predicting the daily precision of the JKII prices in Indonesia's Islamic stock market?" Accordingly, the importance of this research is to deliver efficient techniques in financial stock modeling for financial scholars and decision-makers in the global Islamic stock markets to manage the optimal portfolios and to make the best trading decisions for Islamic stock indices.

The remaining of this research is organized as follows: the literature review and the research gap are discussed in Section 2 "Literature review". Sample and data, analysis techniques using data mining are addressed in Section 3 "Research methodology". The experimental findings are then presented and discussed in Section 4 "Experimental outputs and discussion". Finally, the conclusion and explanations with practical implications are deliberated in Section 5 "Conclusion".

2. Literature review

2.1 Theories and classical analysis of capital market movements

The analysis of capital market movements witnessed a lot of developments, from the use of classical models to modern statistical techniques (Ruhani *et al.*, 2018). The first use of classical models starts with the financial theory on capital markets movements and portfolio risk, volatility and return analyses. It was established by Markowitz (1952) and had supporters like Tobin (1958) and Sharpe (1964), but its significance was not realized until about the 1970s.

Markowitz (1952) developed the modern portfolio theory, which determines the portfolio's risk as systematic (risk at the macroeconomic level) and nonsystematic (microlevel risk). Markowitz (1952) stated that by integrating all assets and all risk, the portfolios with the highest required return degree of risk may be determined from a collection of all potential portfolios. This was referred to as the "Efficient Frontier".

In 1958, Tobin (1958) developed the separation theorem, which argues that the two investment options created by individuals are distinct and separable. As well, Modigliani and Miller (1958) proposed the theory of capital structure, which asserts that, if capital markets are efficient and stabilized, the value of a corporate represents the sum of the market prices of its debt is independent of the amount and form of the debt and equity financing.

In 1964, Sharpe (1964) established the "Capital Asset Pricing Model (CAPM)", which is based on Markowitz (1952) and Tobin's fundamental work (1958). The CAPM states that the location of the tangent line between the capital market line and the efficient frontier is where the super-efficient portfolio formed by combining risk-free and risky assets is situated.

In 1970, Fama (1970) established the efficient capital market theory (also recognized as the random walk theory), which attempts to address the volatility of stock prices. According to Fama (1970), an efficient capital market is one in which price levels always represent all available information.

In 1973, Black and Scholes (1973) developed the "Black-Scholes Formula" founded on the idea that traders can effectively change their portfolios and predict the volatility of the underlying stock. In practice, the formula has shown to be extremely applicable. It is now used globally in every derivatives market.

In addition, Ross (1976) established the arbitrage pricing theory, which became the primary analytical instrument for understanding the price movements of capital markets. This theory is an alternate asset-pricing approach to the CAPM that differs in its fundamental assumptions and analysis of the risk variables correlated with the investment's volatility.

During the 1990s, the academic debate turned from econometric studies of time series on capital market prices, volatilities and profits to the development of methods of human psychology as it pertains to financial markets, culminating in the emergence of behavioral finance (Shiller, 2003). Behavioral finance seeks to clarify and enhance comprehension of the thinking patterns of traders, including their emotional responses and the depth to which these impact their decision-making (Shiller, 2003).

2.2 Modern analysis techniques of capital market movements

According to Prasad and Seetharaman (2021), with information technology and advancement in computing skills, traders began to apply computer programing skills to optimize trading strategies. Sale and purchase in algorithmic trading happen throughout a code such as Python once backtesting requirements are met. In general, following stock prices is challenging as they do not maintain a regular trend and change behaviors often. A present plan might quickly become outdated. As a result, there must be a need for study in the trading industry to profit from the stock market. Researchers in computer science and engineering have recently become interested in using AI to produce trading signals.

Lately, applying data mining techniques is witnessing a large use in many empirical studies due to its effectiveness in predicting prices and volatilities in both Islamic and conventional stock markets, predicting risk and soundness in banks and financial firms, and also in predicting economic events like oil shocks, exchange rate volatilities and other related subjects (Busari *et al.*, 2021; Ledhem, 2021).

By focusing on the adopted technique of AdaBoost in the predicting process of volatilities, trends, distress, returns, risk and prices in stock markets and financial firms, Busari *et al.* (2021) predicted the prices of the Korea Composite Stock Index using the AdaBoost technique. His empirical findings showed that AdaBoost provides high predicting precision compared to other predicting techniques like long short-term memory (LSTM) neural networks and gated recurrent unit neural networks. In the same line, Ampomah *et al.* (2021) applied the AdaBoost technique for predicting the movement of several indices in the stock markets of NASDAQ, NYSE and NSE. Their results demonstrated that the AdaBoost technique provided high accuracy for predicting the movement of stock markets. In another similar study, Zhu *et al.* (2021) predicted the closing price over 180 stock indices using the AdaBoost technique, Elamn model and artificial fish swarm process. Their experimental results showed that the AdaBoost technique outperforms other applied techniques in reducing prediction errors. Correspondingly, Singh (2022), Tunio *et al.* (2021), Zhang *et al.* (2016) and Sun *et al.* (2011) indicated that the AdaBoost technique is crucial for the predicting process with high accuracy based on the lowest predicting errors among other used predicting methods. In summary, all these studies have agreed on the effectiveness of the AdaBoost technique in the predicting process of stock market indices. However, this technique was not used in the Islamic stock market like in Indonesia's Islamic stock market. Therefore, this study is using the AdaBoost technique as an effective data mining technique in predicting the daily precision improvement of JKII prices in Indonesia's Islamic stock market.

In addition, many studies on predicting prices, volatilities and trends in stock markets are performed using a robust data mining technique, which is the K-nearest neighbor (KNN) technique. In a recent study, Khattak *et al.* (2022) predicted the trend of the KSE-100 index in the Pakistan stock exchange using the KNN technique. Their experimental outputs indicated that the KNN technique generated a high level of accuracy in limiting the errors in the prediction process. In the same vein, Latha *et al.* (2022) predicted the movement of several stock market indices using the KNN technique. Their experimental findings indicated that the KNN technique is highly robust for the prediction process of stock movements without numeric prediction problems. Similarly, Tanuwijaya and Hansun (2019), Chen and Hao (2017), Khedr and Yaseen (2017) and Nayak *et al.* (2015) showed that the KNN technique generates high prediction capability and it can be adapted to other exchange market predictions. In summary, these mentioned studies have approved the efficiency of the KNN technique in the predicting process of stock market indices. Nevertheless, this data mining technique was not applied in the cases of the Islamic stock markets. Therefore, this research is applying the KNN technique as an effective data mining technique in predicting the daily precision improvement of JKII prices in Indonesia's Islamic stock market.

As well, different studies on stock market prediction of prices, volatilities and trends in stock marks are conducted using one of the most efficient data mining techniques, which is the random forest (RF) technique. In a current study, Park *et al.* (2022) predicted the returns of S&P 500, KOSPI200 and SSE stock markets indices using the RF technique and LSTM neural network. Their experimental results revealed that the RF technique delivers high predicting accuracy for predicting the returns in the stock exchange markets. In another recent study, Ghosh *et al.* (2022) predicted the directional movements of the S&P 500 stock index using the RF technique and the LSTM neural network. Their experimental findings indicated that the RF technique is highly recommended for predicting the movements of the stock markets index. Similarly, Abraham *et al.* (2022), Tan *et al.* (2019), Vijh *et al.* (2020) and Lohrmann and Luukka (2019) applied the RF technique for classifying and predicting stock index trends. Their findings exposed that the RF technique is highly recommended for the classification and prediction of stock market trends. In summary, studies that have applied the RF technique have agreed on the efficiency of this data mining technique in the predicting process of stock market indices. Yet, this data mining technique was not used in the Islamic stock market like in Indonesia's Islamic stock market. Consequently, this paper is also using the RF technique as an operative data mining technique in predicting the daily precision improvement of JKII prices in Indonesia's Islamic stock market.

In addition, the most common data mining technique for predicting prices, volatilities and trends in stock markets is ANNs (Chhajer *et al.*, 2022). In a recent study, Liu *et al.* (2022) predicted the price trends of Chinese stock markets using ANNs. Their findings indicated that ANNs are useful for predicting stock market trends. Also, Shahvaroughi Farahani and Razavi Hajiagha (2021) predicted the stock market prices of five global stock market indices using the ANN technique and metaheuristic algorithms. Their experimental findings showed that ANN is highly accurate in predicting global stock market prices. Similarly, Wang *et al.* (2020), Livieris *et al.* (2019) and Hu *et al.* (2018) predicted the stock market movements of the S&P 500 index and DJIA index using ANNs. Their experimental outcomes showed that the ANNs technique delivers high prediction performance. Despite the existence of numerous studies on the prediction of conventional stock markets, to the best of the authors' knowledge, there is a limited number of studies that focus on the prediction of the Islamic stock market using ANNs. In a recent study, Ledhem (2022) predicted the prices of the JKII using ANNs. His findings indicated that ANNs are highly suitable for predicting stock market prices. Similarly, Chkili and Hamdi (2021), Aslam *et al.* (2021) and Irsalinda *et al.* (2020) indicated that the ANNs generate accurate prediction performance in limiting the prediction errors of the prediction process in the Islamic capital markets. Thus, the use of ANNs in the predicting process of stock markets is efficient for both Islamic and conventional ones. However, the use of ANNs is not widely spread in the existing studies on predicting Islamic stock markets. Therefore, this study is enriching the existing studies on this subject by using the ANNs as a competent data mining technique for predicting the daily precision improvement of JKII prices in Indonesia's Islamic stock market.

Predicting stock market volatility and prices is becoming increasingly vital in global stock markets. For this purpose, in the past decade, data mining has been extensively used in the prediction process of stock markets. However, in the Islamic stock markets, data mining is not commonly used in predicting the prices and volatilities of Islamic stock indexes. As per the authors' knowledge, no data mining technique other than the ANNs technique was used in the prediction process of Islamic stock markets. This absence of applying data mining methods in the prediction process of Islamic stock markets generates a literature gap. Therefore, by focusing on the target of predicting daily precision improvement of JKII prices in Indonesia's Islamic stock markets using data mining, this study fills this literature gap by applying the AdaBoost, KNN and RF techniques besides the

ANNs as four data mining techniques to provide new efficient techniques for the predicting process in Islamic stock markets.

In addition, one of the newest technologies for improving prediction performance is big data mining (Yang, 2021). Yang *et al.* (2021) used big data mining for risk predicting diabetes. As well, Zhao *et al.* (2022) applied big data mining for evaluating and predicting power marketing. Also, Li (2022) applied big data mining in analyzing sports economic management. Besides, Yue *et al.* (2021) applied big data mining for assessing the financial risk within enterprises. These mentioned studies settled on the usefulness of big data mining, which enhances the prediction performance. For this reason, this study is using big data mining of JKII for achieving the optimal prediction performance of JKII prices.

Consequently, as per the authors' knowledge, the novelty of this paper is to use big data mining of symmetric volatility in the prediction of an Islamic stock market index, which is the JKII. As well, to contribute to the current literature on predicting Islamic stock market indices, this study uses big data mining techniques that have not been used in previous research. Thus, this research fills a gap in the literature and offers new effective techniques for predicting Islamic stock market indices. As a result, it is expected that this study will make a significant addition to the related literature.

3. Research methodology

The implementation of this research for predicting the daily precision improvement of JKII prices is established under four stages. The first stage includes the collection of big data with symmetric volatility of JKII prices. The second stage includes big data evaluation and processing. Then, the third stage includes the application of four data mining techniques using the evaluated big data of JKII from the second stage. Finally, the fourth stage includes the training and testing process, which settles on choosing the optimal big data mining technique method according to the optimal performance score of prediction using four evaluation metrics.

3.1 Big data collection

Following the research of Singh and Yassine (2018), which used big data to predict stock market trends due to the abundance of stock index volatility information that only big data can offer, this research uses a big data sample of symmetric volatility for the daily volatility of JKII. Big data of symmetric volatility were collected from the Investing database which includes the CP, LP, HP, OP and daily JKII actual prices starting from October 6, 2000 to March 30, 2022. Because choosing this period is optimal for collecting the big balance data sheet of JKII's financial data from Investing database to use it for extracting the largest possible hidden knowledge from it, which helps to predict daily precision improvement of JKII prices using data mining. The four attributes of the symmetric volatility framework are used as inputs (CP, LP, HP and OP) based on studies on predicting stock markets indices by Nayak *et al.* (2015), Khedr and Yaseen (2017), Vijh *et al.* (2020), Ampomah *et al.* (2021) and Ledhem (2022). Although the output layer comprises a single target to be predicted, which is the daily actual JKII prices.

3.1.1 Big data evaluation and processing. In the big data set, there are 24,970 observations on-balance sheet items with no missing values or outliers, which indicates that the data was in a suitable form for the analysis and modeling with high accuracy and effectiveness of data mining and low-biased results. This big data sample is divided into two subsets: training data set and testing data set. The training process took 8 h with 10 procedures for repeated training and testing with a training set size of 75%.

3.2 Data mining techniques methodology

To accomplish optimal prediction performance for the JKII prices, this research applies several big data mining techniques using the data science of Python like AdaBoost, KNN, RF and ANNs.

3.2.1 AdaBoost technique. The AdaBoost technique (also known as Adaptive Boosting) is a relatively new data mining algorithm developed by Freund and Schapire (1997). According to Sun *et al.* (2011), the AdaBoost algorithm process is as follows:

Supposing that $T_n = \{(x_1, y_1),\ldots, (x_n, y_n)\}$ is the training set samples with $y_i(i = 1, 2,\ldots,n)$ $\in\{-1,1\}$, in which y_i is the vector of the observed variable being predicted as output, whereas x_i is the vector of input variables, it signifies two classes for simplification determination.

The weight distribution over the set samples at the t^{th} iteration of boosting is signified as:

$$W_t = \left\{w_t^1, w_t^2, \ldots, w_t^n\right\} \quad (t = 1, 2, \ldots, T) \tag{1}$$

This weight distribution is set in a uniform way from the beginning, in which w_t^i $(i = 1, 2, \ldots, n)$ holds $1/n$ value at the beginning of iteration when $t = 1$ with an adaptive update at the remaining iterations. In each t iteration, the AdaBoost establishes a new training data set by resampling the starting iteration data set with W_t (weight distribution) over the algorithm of weak learner for building a base classifier signified as C_t in this new training data set.

Then, C_t is applied to classify samples in the first data set, where the error E_t of C_t is calculated as the following:

$$E_t = \sum\nolimits_{i:C_t(x_i)\neq y_i} w_t^i \tag{2}$$

After that, the weight distribution of W_t samples are updated as the following:

$$W'^i_{t+1} = W_t^i \cdot \exp\left(-\alpha_t \cdot l_t^i\right)(i = 1, 2, \ldots, n) \tag{3}$$

in which α_t and l_t^i are calculated as the following:

$$\alpha_t = 0,5\ln\left(\frac{1 - E_t}{E_t}\right) \tag{4}$$

$$l_t^i = \begin{cases} 1 \ if & C_t(x_i) = \ y_i \\ -1 \ if & C_t(x_1) \neq \ y_i \end{cases} \tag{5}$$

The above-updated weights are then normalized as follows:

$$W_{t+1}^i = \frac{W'^i_{t+1}}{\sum_{i=1}^n W'^i_{t+1}} \ (i = 1, 2, \ldots, n) \tag{6}$$

For each t iteration is processed, the ensemble of $C = \{C_1, C_2,\ldots,C_T\}$ is composed within the T weak classifiers. Finally, the final classification of AdaBoost output is completed through a set of their classification outputs weighted by $A = \{\alpha_1, \alpha_2,\ldots,\alpha_t\}$.

In this study, the number of estimator weights is applied under 50 estimator weights within the classification boosting algorithm of SAMME.R developed by Hastie *et al.* (2009).

3.2.2 K-nearest neighbor technique. As stated by Ledhem (2021), Urso *et al.* (2019) and Hastie *et al.* (2005), KNN serves as a regression approach when implemented in a data set of cases (x_i; y_i) specified by numeric characteristics as an aspect of the vector x_i and corresponding numeric value y_i. As a result, it is important to assess the value $\hat{y}$. In accordance with a hidden occurrence vector X based on KNN knowledge about the training process, which deliberates the simple occurrence of the data set used by a specific attribute. The i^{th} case of the data set is defined by the pair (x_i; y_i), with x_i being a scalar.

Given a new instance ($x = k; \hat{y} = ?$), KNN runs an output $\hat{y}$ calculated based on the following equation:

$$\hat{y} = \frac{1}{k} \sum\nolimits_{x_i \in N_k(x)} y_i \tag{7}$$

where $N_k(x)$ is the collective ensemble of neighbors $\{x_i, \ldots, x_k\}$ from the training operation of a new value x and $N_k(x)$ persevere on the selected value for K.

For the prediction role, KNN does not use any clear prediction form, which makes the KNN regression a nonparametric approach.

Following the predicting process of financial performance of Indonesian Islamic banks by Ledhem (2021) using KNN with the nearest training of instances set based on the cross-validation method, this research applies the cross-validation method, which involves dividing the data set into several parts and training the model on various combinations of these partitions. By comparing the model's performance with diverse values of K, this research determines that the optimal performance was attained using 5 neighbors with uniform weight and Euclidean metric for the training process.

3.2.3 Random forest technique. RF is a collection of decision trees that are built using the most successful method for increasing tree variety(Granitto *et al.*, 2007). Decision trees are quite unstable because they are techniques where even a small change in the data set may generate significant changes in the created model (Breiman, 1996). Thus, to increase the variety of cluster memberships and reduce unwanted changes, RF uses a bootstrap procedure to adapt each tree to the overall sample set. A random subset of an existing data set with the same length, reserved with supplementary data, is used in the bootstrap procedure (Efron and Tibshirani, 1994).

During the development of every tree, another variety mechanism is presented, because on each node, a process in RF selects a tiny random subclass of features and searches for the optimal split using just this subclass (Terissi *et al.*, 2018). The combination of the two variety mechanisms (bootstrapping with selecting across every single node just from a subclass of features) results in improved ensemble construction and resilient performance for a data mining predicting approach (Granitto *et al.*, 2007; Terissi *et al.*, 2018).

This research uses these two mechanisms in RF with 10 trees based on categorized data at the 95% confidence level.

3.2.4 Artificial neural networks technique. Stock market behavior is regarded to be wholly random and substantially nonlinear, according to Vachhani *et al.* (2020). The ANN's capacity to retain specific data for subsequent use makes it the most efficient way for enhancing the ability to analyze the complex structure between interconnections of stock price data.

ANNs are the most widely used data mining algorithms for both predicting and categorization. The biological neurons network design inspires the ANN method, in which neurons are interconnected and emulated in experiential learning. An ANN is made up of layers of neurons that are completely linked to the preceding layer through a weighting scheme. Several distinct neural structures have been planned in advance. The most successful uses are multilayer feed-forward neural networks, which have been shown in related studies of capital markets prediction. ANNs of this type are made up of an input layer, one or much more hidden layers, and the output layer (Shmueli *et al.*, 2017).

The input layer contains the input parameters that are supplied and transmitted to the hidden layer, which is then passed on to the output layer (Giudici and Figini, 2009; Shmueli *et al.*, 2017).

The output of the hidden layer nodes is computed as follows. A weighted total of inputs is determined first and then a technique called the transference function is conducted to this summation (Shmueli *et al.*, 2017).

Supposing the input values are $x_1, x_2, \ldots, x_m$. The node output j is derived using the weighted sum $\theta_J + \sum_{i=1}^{m} \omega_{ij} x_{ij}$, knowing that θ_J, $\omega_{1j}, \ldots, \omega_{mj}$ are weights for largely randomized created and controlled sets as the network learns. The transference function is then applied to the mentioned sum, with the monotone function serving as the transference function. Although the logistic function A is the most frequent transference for the role of activation and it is determined as follows, according to Shmueli *et al.* (2017):

$$A(s) = \frac{1}{(1 + e^{-s})} \tag{8}$$

Finally, the output layer receives information from the input layer through the hidden layer, with an iterated process based on the same transference function, which is used for producing the output (Shmueli *et al.*, 2017).

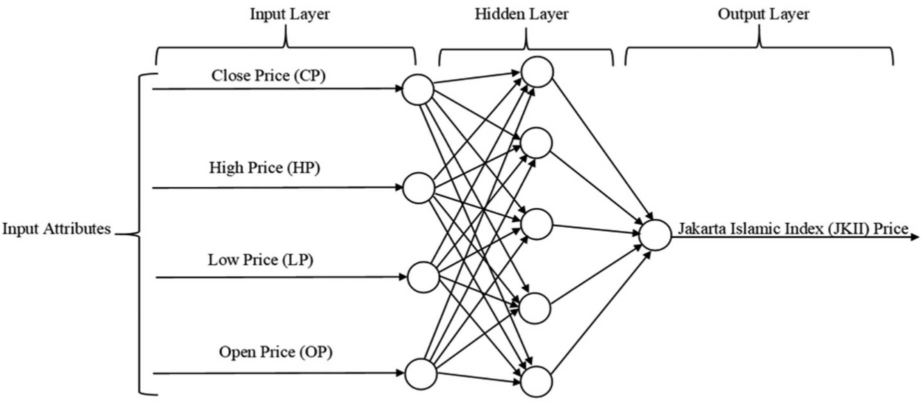
In the ANNs technique, the back-propagation algorithm is the most widely used method for weights activation and updating (Rumelhart *et al.*, 1988). Nonetheless, this typical method has a slow learning rate (Castillo *et al.*, 2006). As a result, the algorithm of Broyden–Fletcher–Goldfarb–Shanno for activating weights is used in this research to solve this problem, with the logistic function for activating ANN, 2,000 neurons in the hidden layer, alpha regularization = 0.0001, and 100 iterations, shows the ANN structure for predicting daily precision improvement of JKII prices in Indonesia's Islamic stock market is shown in Figure 1.

4. Experimental outputs and discussion

The main operation in this research before the prediction process is choosing the optimal big data mining technique method according to the optimal performance score of training and testing (Figure 2).

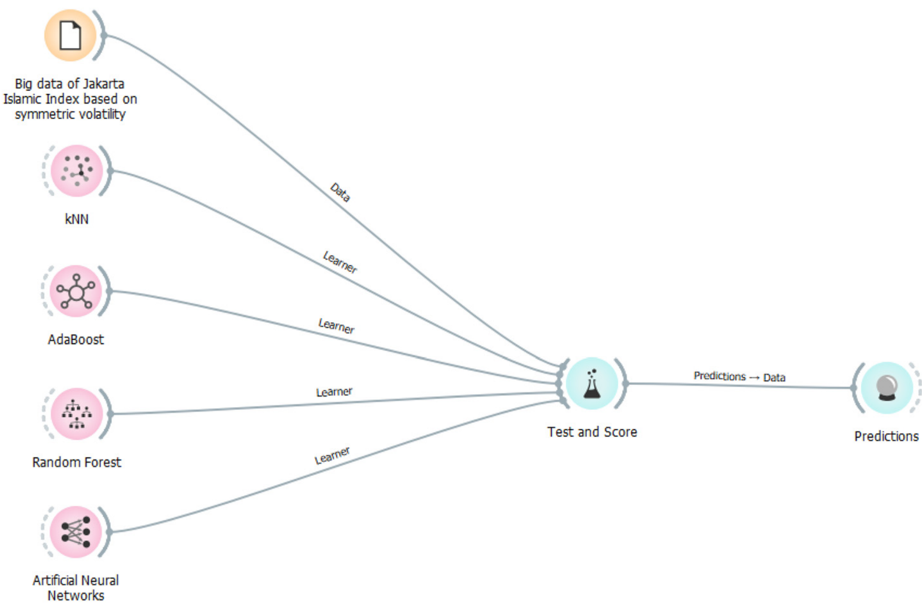
4.1 Training and testing

Following studies by Busari *et al.* (2021), Ledhem (2021), Singh (2022), Ying *et al.* (2021) and Zhu *et al.* (2021) that use the standard metrics for choosing the optimal data mining technique, this study uses four metrics (mean absolute error [MAE], mean squared error [MSE] and root mean squared error [RMSE] and R-squared) standard metrics to reveal the optimal performance for each used data mining technique. Pham *et al.* (2020) state that the



Source: Prepared by authors

Figure 1.
ANN structure for predicting daily precision improvement of JKII prices in Indonesia's Islamic stock market



Source: Summary of predicting daily precision improvement of JKII prices using data science of Python

Figure 2.
Summary of predicting daily precision improvement of JKII prices in Indonesia's Islamic stock market by data mining techniques

RMSE is calculated using the MSE or estimate of variance unexplained by the research characteristics. While MSE is essential for calculating the *R*-squared, or maximum coefficient of determination, for a model of systemic impact levels. The *R*-squared number is the percentage of variation that the model successfully explains. One estimation of the efficiency limit for a predicted JKII pricing is provided by the *R*-squared. While MAE is a more reasonable way of measuring average error and (unlike RMSE) is unequivocal.

Willmott and Matsuura (2005) proved that MAE should serve as a baseline for dimensional assessments and intercomparisons of average model-efficiency error.

According to Chicco *et al.* (2021), these metrics are defined as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i\right)^2 \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i\right)^2} \tag{10}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i - \hat{Y}_i\right| \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(X_i - Y_i\right)^2}{\sum_{i=1}^{N}\left(\overline{Y} - Y_i\right)^2}, \tag{12}$$

in which N signifies the input number. While X_i signifies the inputs values, Y_i and $\hat{Y}_i$ signifies the actual value and the predicted target respectively. $\overline{Y}$ is the mean of true values that are defined as follows:

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \tag{13}$$

The lower values of MSE, RMSE and MAE metrics signify higher prediction performance in which the closest value to 0 signifies the highest prediction performance. While closer values of R-squared to 1 signify higher prediction performance.

The evaluation findings of training and testing demonstrate that the optimal performance for predicting daily precision improvement of JKII prices in Indonesia's Islamic stock market is by applying the AdaBoost technique with an MSE = 0.0001196, RMSE = 0.010936, MAE = 0.0024089 and R^2 = 0.999999. In addition, the RF technique delivers closer values to the optimal performance of the AdaBoost with an MSE = 0.02473, RMSE = 0.157271, MAE = 0.081949 and R^2 = 0.9999995 (Table 1).

4.2 Outputs discussion and visualization

This research performs four big data mining techniques to predict daily precision improvement of JKII prices for seven anticipated days starting from March 31, 2022 until April 08, 2022 in Indonesia's Islamic stock market (Table A1) as is shown in Figures 3–6. The experimental findings indicated that both the AdaBoost technique and RF are the most suitable techniques for the predicting process. This prediction efficiency can be seen in the congruence of the predicted curve over the curve of actual values in Figures 3 and 4.

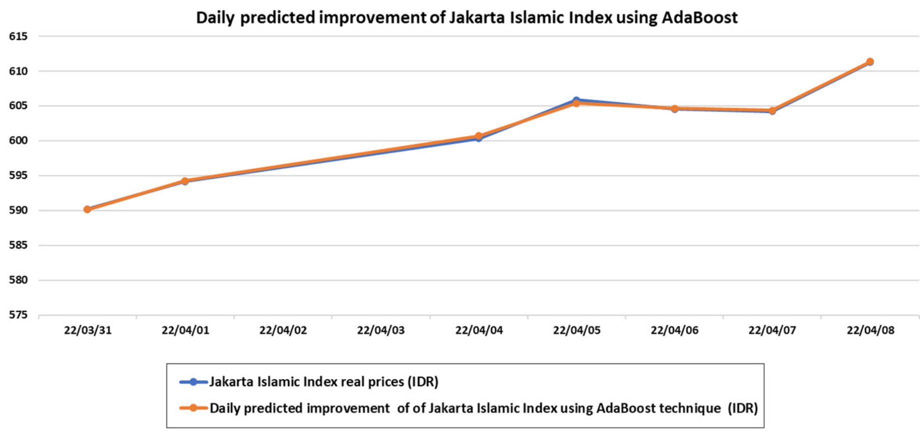
However, as the AdaBoost technique provides the best precision performance for prediction, it is better to perform the AdaBoost for predicting daily precision improvement of JKII prices in Indonesia's Islamic stock market. Whereas, by using the KNN, the curve of predicted values is not matching the curve of actual values in Figure 5, but does not move away from it for many days. As for the use of ANNs, it is seen that the prediction curve moves away from the curve of the actual values in many days as shown in Figure 6, which indicates the low prediction performance of this technique compared to other used techniques.

To check and validate the robustness of the AdaBoost technique in the prediction process alongside other used techniques, this research compares the predicted JKII prices with the matching actual values by measuring absolute values of differences between actual JKII' daily prices and the predicted ones (Δ = Actual prices of JKII − Predicted prices of JKII) in the seven predicted days. For this reason, Table 2 shows that the lowest absolute difference refers to the AdaBoost technique in most of the predicting days except for two cases (April 01, 2022 and April 07, 2022) when these lowest differences refer to the RF technique. These results validate the high rate of predicting performance in the case of applying the AdaBoost technique for predicting daily precision improvement of JKII prices in Indonesia's Islamic stock markets as shown in Figure 7. Figure 7 shows that absolute values Δ of the differences between the actual and predicted values of JKII prices using the AdaBoost technique are concentrated in the closest area to 0 compared to other data mining

Technique	MSE	RMSE	MAE	R^2
AdaBoost	0.000119611*	0.0109366*	0.0024089*	0.9999999*
RF	0.02473417	0.15727102	0.08194963	0.9999995
KNN	0.951410	0.9754026	0.595086335	0.9999823
ANNs	1.5690475	1.25261628	0.945094331	0.99997092

Note: *Refers to the optimal prediction performance
Source: Prediction performance outcomes of testing and training using data science of Python

Table 1. Assessment outcomes of testing and training



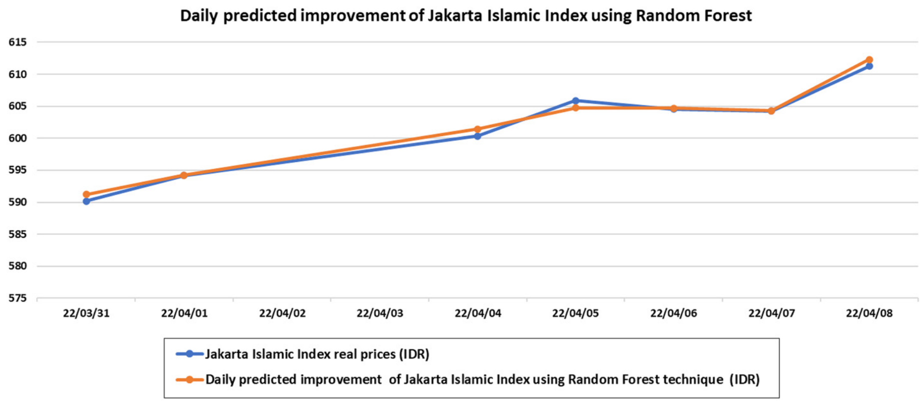
Source: Organized by authors

Figure 3. Predictions of daily precision improvement of JKII prices in Indonesia's Islamic stock market using the AdaBoost technique

techniques that diverge from 0, which indicates that the predicted prices of JKII using this technique are closer to the actual JKII prices compared to other used techniques. This result is reliable to the findings of Liu *et al.* (2021) in which the AdaBoost technique provides unique advantages compared to other data mining techniques in terms of the prediction accuracy of capital markets movements. Also, AdaBoost generates low error prediction values and its algorithmic analysis coding is easy and does not require adjustments of the input parameters as stated by Liu *et al.* (2021).

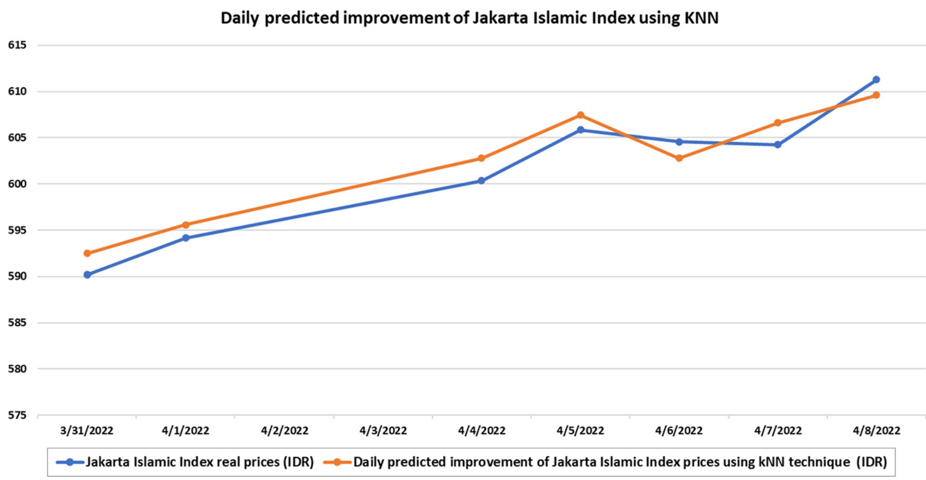
5. Conclusion

This research focuses mostly on making accurate predictions about the Indonesian financial market. Stock price prediction is primarily done to reduce the significant losses experienced



Source: Organized by authors

Figure 4.
Predictions of daily precision improvement of JKII prices in Indonesia's Islamic stock market using the RF technique



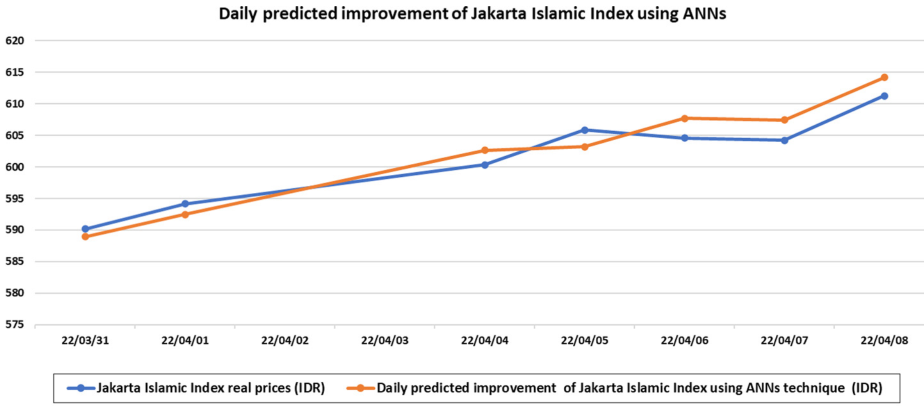
Source: Organized by authors

Figure 5.
Predictions of daily precision improvement of JKII prices in Indonesia's Islamic stock market using the KNN technique

by investors and to analyze the best investment choices. Thus, this research aims to predict daily precision improvement of JKII prices based on big data mining of symmetric volatility framework. Due to the absence of using big data mining techniques in the predicting process of Islamic stock markets, this research contributes to the existing studies by providing new efficient techniques for predicting Islamic stock market indices. This research applies four techniques of data mining (AdaBoost, RF, KNN and ANNs) and selects the optimal technique based on the optimal performance of the training and testing process.

The experimental findings proved that AdaBoost and RF are the best techniques for predicting daily precision improvement of JKII prices in Indonesia's Islamic stock market due to the optimal predicting performance according to the criteria of MAE, MSE, RMSE and *R*-squared. However, although the RF provides a closer predicting accuracy to the AdaBoost, the AdaBoost outperforms better than the RF. Therefore, as an answer to the main research question, AdaBoost is the optimal big data mining technique for predicting daily precision improvement of JKII prices in Indonesia's Islamic stock market.

To check and validate the robustness of the AdaBoost technique in the prediction process alongside other used techniques, this research compares the predicted JKII prices with the matching actual values by measuring the absolute values of the differences between actual JKII's daily prices and the predicted ones in the seven predicted days. The



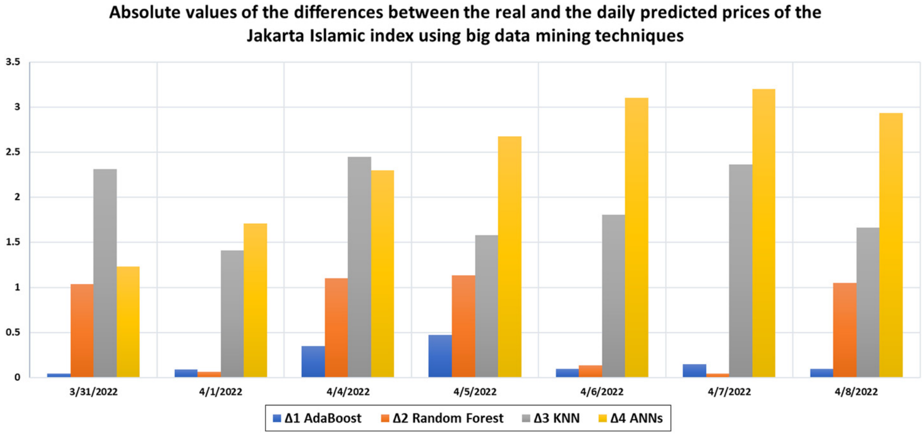
Source: Organized by authors

Figure 6. Predictions of daily precision improvement of JKII prices in Indonesia's Islamic stock market using ANNs technique

Days	Δ1 AdaBoost	Δ2 Random forest	Δ3 KNN	Δ4 ANNs
March 31, 2022	0.046*	1.036	2.314	1.235
April 01, 2022	0.092	0.063*	1.410	1.707
April 04, 2022	0.350*	1.104	2.448	2.299
April 05, 2022	0.474*	1.134	1.576	2.675
April 06, 2022	0.096*	0.138	1.804	3.103
April 07, 2022	0.149	0.045*	2.362	3.198
April 08, 2022	0.096*	1.050	1.662	2.935

Note: Δ = |Actual JKII − Predicted JK|. *Lowest difference between the actual and predicted JKII prices
Source: Prepared by authors

Table 2. Comparison between actual daily JKII prices and predicted daily JKII prices



Source: Organized by authors

Figure 7. Concentration of the absolute values of the differences between the actual and predicted values of JKII prices

experimental results validate the high precision of predicting performance, which refers to the AdaBoost technique with the lowest absolute differences in most of the predicting days of JKII prices.

In addition, to check the prediction performance of the AdaBoost technique, this study performs a final verification according to the actual-world predicting scenario, by calculating the absolute values of the differences between actual JKII's daily prices and the predicted ones. The results demonstrate the robustness of predicting performance when using the AdaBoost technique, which has the prospect to be a valuable tool for financial analyzers, financial scholars and investors by providing them with the necessary knowledge to achieve accurate portfolio management, minimizing trading risk and make wise financial choices in the Islamic stock markets.

Above and beyond, Islamic stock markets still require modern techniques such as data mining techniques, which are useful methods in the stock markets and can offer extensive knowledge to the investors for gaining wise management of portfolios by decreasing trading uncertainty and selecting the most suitable investment decisions, which results to the certainty due to the efficiency of those advanced techniques in the predicting process than classical statistical methods that are paralyzed to analyze big data in the stock market. Therefore, this fact is obliging all financial analyzers, financial scholars and investors to apply data mining techniques that outperform the classical statistical methods in analyzing, classifying and predicting global Islamic stock market indices.

5.1 Practical implications for research and practice

- This research aims to predict daily precision improvement for the JKII prices using big data mining of symmetric volatility. Consequently, this research is filling a literature gap by presenting effective techniques for predicting movements, volatilities and prices in the global Islamic stock markets by helping investors to manage risk.
- This research informs investment strategies in Islamic stock markets by providing insights on how to optimize portfolios based on predicting daily precision

improvement of capital market indices. As well, it provides market analysis insights that can inform market analysis and strategic decision-making.

- Although this research is limited to the Islamic stock index which is JKII, the used big data mining techniques can be operative techniques for financial analyzers, financial scholars and investors to estimate and predict the volatilities and uncertainties in both Islamic and standard stock markets. The study can be extended to other Islamic stock markets to provide a comparative analysis of market movements and trends.

5.2 Limitations and future research

To predict the JKII daily precision improvement using big data mining, this research restricted the prediction modeling through the symmetric volatility information. Future research could consider using asymmetric volatility information for predicting the JKII daily precision improvement. Also, this research can be extended to other global Islamic stock markets to provide a comparative analysis of market movements and trends. As well, future research can focus on developing more sophisticated portfolio management strategies for global Islamic capital markets that incorporate ML, DL and other big data mining techniques.

In addition, although data mining techniques improve the prediction accuracy of portfolio management, data mining techniques may not essentially take into account the multiple criteria that decision-makers consider when assessing investment options. This is where multi-criteria decision-making (MCDM) techniques can complement data mining techniques to improve investment portfolio decisions in capital markets (Basilio *et al.*, 2018; Sachdeva *et al.*, 2021; Yeşildağ *et al.*, 2020). Thus, future research could consider adding MCDM techniques to big data mining for optimizing investment decisions in Islamic capital markets by predicting Islamic stock movements and risks.

References

Abraham, R., Samad, M.E., Bakhach, A.M., El-Chaarani, H., Sardouk, A., Nemar, S.E. and Jaber, D. (2022), "Forecasting a stock trend using genetic algorithm and random forest", *Journal of Risk and Financial Management*, Vol. 15 No. 5, p. 188.

Albaity, M. and Ahmad, R. (2008), "Performance of Syariah and composite indices: evidence from bursa Malaysia", *Asian Academy of Management Journal of Accounting and Finance*, Vol. 4 No. 1, pp. 23-43.

Al-Zoubi, H.A. and Maghyereh, A.I. (2007), "The relative risk performance of Islamic finance: a new guide to less risky investments", *International Journal of Theoretical and Applied Finance*, Vol. 10 No. 2, pp. 235-249.

Ampomah, E.K., Qin, Z., Nyame, G. and Botchey, F.E. (2021), "Stock market decision support modeling with tree-based AdaBoost ensemble machine learning models", *Informatica*, Vol. 44 No. 4, doi: 10.31449/inf.v44i4.3159.

Aslam, F., Mughal, K.S., Ali, A. and Mohmand, Y.T. (2021), "Forecasting Islamic securities index using artificial neural networks: performance evaluation of technical indicators", *Journal of Economic and Administrative Sciences*, Vol. 37 No. 2, pp. 253-271.

Basilio, M.P., de Freitas, J.G., Kämpffe, M.G.F. and Bordeaux Rego, R. (2018), "Investment portfolio formation via multicriteria decision aid: a Brazilian stock market study", *Journal of Modelling in Management*, Vol. 13 No. 2, pp. 394-417, doi: 10.1108/JM2-02-2017-0021.

Ben Rejeb, A. and Arfaoui, M. (2019), "Do Islamic stock indexes outperform conventional stock indexes? A state space modeling approach", *European Journal of Management and Business Economics*, Vol. 28 No. 3, pp. 301-322.

Black, F. and Scholes, M. (1973), "The pricing of options and corporate liabilities", *Journal of Political Economy*, Vol. 81 No. 3, pp. 637-654.

Breiman, L. (1996), "Bagging predictors", *Machine Learning*, Vol. 24 No. 2, pp. 123-140.

Busari, G.A., Kwak, N.W. and Lim, D.H. (2021), "An application of AdaBoost-GRU ensemble model to economic time series prediction", *Indian Journal of Science and Technology*, Vol. 14 No. 31, pp. 2557-2566.

Castillo, E., Guijarro-Berdiñas, B., Fontenla-Romero, O. and Alonso-Betanzos, A. (2006), "A very fast learning method for neural networks based on sensitivity analysis", *Journal of Machine Learning Research*, Vol. 7, pp. 1159-1182.

Chen, Y. and Hao, Y. (2017), "A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction", *Expert Systems with Applications*, Vol. 80, pp. 340-355.

Chhajer, P., Shah, M. and Kshirsagar, A. (2022), "The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction", *Decision Analytics Journal*, Vol. 2, p. 100015.

Chicco, D., Warrens, M.J. and Jurman, G. (2021), "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation", *PeerJ Computer Science*, Vol. 7, p. e623.

Chkili, W. and Hamdi, M. (2021), "An artificial neural network augmented GARCH model for Islamic stock market volatility: do asymmetry and long memory matter?" *International Journal of Islamic and Middle Eastern Finance and Management*, Vol. 14 No. 5, pp. 853-873.

Efron, B. and Tibshirani, R.J. (1994), *An Introduction to the Bootstrap*, CRC press.

Fama, E.F. (1970), "Efficient capital markets: a review of theory and empirical work", *The Journal of Finance*, Vol. 25 No. 2, pp. 383-417.

Freund, Y. and Schapire, R.E. (1997), "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of Computer and System Sciences*, Vol. 55 No. 1, pp. 119-139.

Ghosh, P., Neufeld, A. and Sahoo, J.K. (2022), "Forecasting directional movements of stock prices for intraday trading using LSTM and random forests", *Finance Research Letters*, Vol. 46, p. 102280.

Giudici, P. and Figini, S. (2009), *Applied Data Mining for Business and Industry*, John Wiley and Sons.

Granitto, P.M., Gasperi, F., Biasioli, F., Trainotti, E. and Furlanello, C. (2007), "Modern data mining tools in descriptive sensory analysis: a case study with a random forest approach", *Food Quality and Preference*, Vol. 18 No. 4, pp. 681-689.

Hastie, T., Rosset, S., Zhu, J. and Zou, H. (2009), "Multi-class AdaBoost", *Statistics and Its Interface*, Vol. 2 No. 3, pp. 349-360.

Hastie, T., Tibshirani, R., Friedman, J. and Franklin, J. (2005), "The elements of statistical learning: data mining, inference and prediction", *The Mathematical Intelligencer*, Vol. 27 No. 2, pp. 83-85.

Hu, H., Tang, L., Zhang, S. and Wang, H. (2018), "Predicting the direction of stock markets using optimized neural networks with Google trends", *Neurocomputing*, Vol. 285, pp. 188-195.

Irsalinda, N., Astuti, Y.S. and Sugiyarto. (2020), "Fuzzy feed forward neural network (FFFNN) model for the Jakarta Islamic index (JII) forecasting", *IOP Conference Series: Materials Science and Engineering*, Vol. 821 No. 1, p. 012004.

Jannes, K. (2018), *Machine Learning for Finance: Data Algorithms for the Markets and Deep Learning from the Ground up for Financial Experts and Economics*, Packt Publishing, Birmingham, UK (accessed 22 May 2021).

Khattak, A., Khan, A., Ullah, H., Asghar, M.U., Arif, A., Kundi, F.M. and Asghar, M.Z. (2022), "An efficient supervised machine learning technique for forecasting stock market trends", *Information and Knowledge in Internet of Things*, Springer, pp. 143-162.

Khedr, A.E. and Yaseen, N. (2017), "Predicting stock market behavior using data mining technique and news sentiment analysis", *International Journal of Intelligent Systems and Applications*, Vol. 9 No. 7, p. 22.

Latha, R.S., Sreekanth, G.R., Suganthe, R.C., Geetha, M., Selvaraj, R.E., Balaji, S., Harini, K.R., *et al.* (2022), "Stock movement prediction using KNN machine learning algorithm", *2022 International Conference on Computer Communication and Informatics (ICCCI)*, *IEEE*, pp. 1-5.

Ledhem, M.A. (2021), "Data mining techniques for predicting the financial performance of Islamic banking in Indonesia", *Journal of Modelling in Management*, Vol. 17 No. 3, pp. 896-915, doi: 10.1108/JM2-10-2020-0286.

Ledhem, M.A. (2022), "Deep learning with small and big data of symmetric volatility information for predicting daily accuracy improvement of JKII prices", *Journal of Capital Markets Studies*, Vol. 6 No. 2, pp. 130-147, doi: 10.1108/JCMS-12-2021-0041.

Li, Q. (2022), "Application of multisource big data mining technology in sports economic management analysis", *Mathematical Problems in Engineering*, Vol. 2022.

Liu, H., Huang, S., Wang, P. and Li, Z. (2021), "A review of data mining methods in financial markets", *Data Science in Finance and Economics*, Vol. 1 No. 4, pp. 362-392.

Liu, Q., Tao, Z., Tse, Y. and Wang, C. (2022), "Stock market prediction with deep learning: the case of China", *Finance Research Letters*, Vol. 46, p. 102209.

Livieris, I.E., Kotsilieris, T., Stavroyiannis, S. and Pintelas, P. (2019), "Forecasting stock price index movement using a constrained deep neural network training algorithm", *Intelligent Decision Technologies*, Vol. 14 No. 3, pp. 313-323.

Lohrmann, C. and Luukka, P. (2019), "Classification of intraday S&P500 returns with a random forest", *International Journal of Forecasting*, Vol. 35 No. 1, pp. 390-407.

Markowitz, H. (1952), "Portfolio selection", *The Journal of Finance, [American Finance Association, Wiley]*, Vol. 7 No. 1, pp. 77-91.

Modigliani, F. and Miller, M.H. (1958), "The cost of capital, corporation finance and the theory of investment", *The American Economic Review*, Vol. 48 No. 3, pp. 261-297.

Nayak, R.K., Mishra, D. and Rath, A.K. (2015), "A naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices", *Applied Soft Computing*, Vol. 35, pp. 670-680.

Othman, A.H.A., Kassim, S., Rosman, R.B. and Redzuan, N.H.B. (2020), "Prediction accuracy improvement for bitcoin market prices based on symmetric volatility information using artificial neural network approach", *Journal of Revenue and Pricing Management*, Vol. 19 No. 5, pp. 314-330.

Park, H.J., Kim, Y. and Kim, H.Y. (2022), "Stock market forecasting using a multi-task approach integrating long short-term memory and the random Forest framework", *Applied Soft Computing*, Vol. 114, p. 108106.

Pham, L.L., Watford, S.M., Pradeep, P., Martin, M.T., Thomas, R.S., Judson, R.S., Setzer, R.W. and Friedman, K.P. (2020), "Variability in in vivo studies: defining the upper limit of performance for predictions of systemic effect levels", *Computational Toxicology*, Vol. 15, p. 100126.

Prasad, A. and Seetharaman, A. (2021), "Importance of machine learning in making investment decision in stock market", *Vikalpa: The Journal for Decision Makers*, Vol. 46 No. 4, pp. 209-222.

Qizam, I., Ardiansyah, M. and Qoyum, A. (2020), "Integration of Islamic capital market in ASEAN-5 countries: preliminary evidence for broader benefits from the post-global financial crisis", *Journal of Islamic Accounting and Business Research*, Vol. 11 No. 4, pp. 811-825.

Ross, S.A. (1976), "The arbitrage theory of capital asset pricing", *Journal of Economic Theory*, Vol. 13 No. 3, pp. 341-360.

Ruhani, F., Islam, M.A.I. and Ahmad, T.S.T. (2018), "Theories explaining stock price behavior: a review of the literature", *International Journal of Islamic Banking and Finance Research*, Vol. 2 No. 2, pp. 51-64.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1988), "Learning representations by back-propagating errors", *Cognitive Modeling*, Vol. 5 No. 3, p. 1.

Sachdeva, M., Lehal, R., Gupta, S. and Garg, A. (2021), "What make investors herd while investing in the Indian stock market? A hybrid approach", *Review of Behavioral Finance*, Vol. 15 No. 1, pp. 19-37, doi: 10.1108/RBF-04-2021-0070.

Shahvaroughi Farahani, M. and Razavi Hajiagha, S.H. (2021), "Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models", *Soft Computing*, Vol. 25 No. 13, pp. 8483-8513.

Sharpe, W.F. (1964), "Capital asset prices: a theory of market equilibrium under conditions of risk", *The Journal of Finance*, Vol. 19 No. 3, pp. 425-442.

Shiller, R.J. (2003), "From efficient markets theory to behavioral finance", *Journal of Economic Perspectives*, Vol. 17 No. 1, pp. 83-104.

Shmueli, G., Bruce, P.C., Yahav, I., Patel, N.R. and Lichtendahl, K.C. Jr (2017), *Data Mining for Business Analytics: Concepts, Techniques, and Applications in R*, John Wiley and Sons.

Singh, G. (2022), "Machine learning models in stock market prediction", *International Journal of Innovative Technology and Exploring Engineering*, Vol. 11 No. 3, pp. 18-28.

Singh, S. and Yassine, A. (2018), "Big data mining of energy time series for behavioral analytics and energy consumption forecasting", *Energies, Multidisciplinary Digital Publishing Institute*, Vol. 11 No. 2, p. 452.

Subekti, R., Abdurakhman. and Rosadi, D. (2020), "Modified capital asset pricing model (CAPM) into sharia framework", *Journal of Physics: Conference Series*, Vol. 1581 No. 1, pp. 1-6.

Sun, J., Jia, M. and Li, H. (2011), "AdaBoost ensemble for financial distress prediction: an empirical comparison with data from Chinese listed companies", *Expert Systems with Applications*, Vol. 38 No. 8, pp. 9305-9312.

Tan, Z., Yan, Z. and Zhu, G. (2019), "Stock selection with random Forest: an exploitation of excess return in the Chinese stock market", *Heliyon*, Vol. 5 No. 8, p. e02310.

Tanuwijaya, J. and Hansun, S. (2019), "LQ45 stock index prediction using k-nearest neighbors regression", *International Journal of Recent Technology and Engineering (IJRTE)*, Vol. 8 No. 3, pp. 2388-2391.

Terissi, L.D., Sad, G.D. and Gómez, J.C. (2018), "Robust front-end for audio, visual and audio–visual speech classification", *International Journal of Speech Technology*, Vol. 21 No. 2, pp. 293-307.

Tobin, J. (1958), "Liquidity preference as behavior towards risk", *The Review of Economic Studies, Oxford University Press, Review of Economic Studies, Ltd*, Vol. 25 No. 2, pp. 65-86.

Tunio, F.H., Ding, Y., Agha, A.N., Agha, K. and Panhwar, H.U.R.Z. (2021), "Financial distress prediction using AdaBoost and bagging in Pakistan stock exchange", *The Journal of Asian Finance, Economics and Business*, Vol. 8 No. 1, pp. 665-673.

Urso, A., Fiannaca, A., La Rosa, M., Ravì, V. and Rizzo, R. (2019), "Data mining: prediction methods", in Ranganathan, S., Gribskov, M., Nakai, K. and Schönbach, C. (Eds), *Encyclopedia of Bioinformatics and Computational Biology*, Academic Press, Oxford, pp. 413-430.

Vachhani, H., Obiadat, M.S., Thakkar, A., Shah, V., Sojitra, R., Bhatia, J. and Tanwar, S. (2020), "Machine learning based stock market analysis: a short survey", in Raj, J.S., Bashar, A. and Ramson, S.R.J. (Eds), *Innovative Data Communication Technologies and Application*, Springer International Publishing, Cham, pp. 12-26, doi: 10.1007/978-3-030-38040-3_2.

Vijh, M., Chandola, D., Tikkiwal, V.A. and Kumar, A. (2020), "Stock closing price prediction using machine learning techniques", *Procedia Computer Science*, Vol. 167, pp. 599-606.

Wang, Y., Wang, L., Yang, F., Di, W. and Chang, Q. (2020), "Advantages of direct input-to-output connections in neural networks: the Elman network for stock index forecasting", *Information Sciences*, Vol. 547, pp. 1066-1079.

Willmott, C.J. and Matsuura, K. (2005), "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", *Climate Research*, Vol. 30 No. 1, pp. 79-82.

Yang, H. (2021), "Application value of big data mining technology in the field of financial venture capital", *International Conference on Applications and Techniques in Cyber Security and Intelligence*, Springer, pp. 333-340.

Yang, H., Luo, Y., Ren, X., Wu, M., He, X., Peng, B., Deng, K., *et al.* (2021), "Risk prediction of diabetes: big data mining with fusion of multifarious physical examination indicators", *Information Fusion*, Vol. 75, pp. 140-149.

Yeşildağ, E., Özen, E. and Baykut, E. (2020), "An assessment for financial performance of banks listed in Borsa Istanbul by multiple criteria decision making", in Özen, E. and Grima, S. (Eds), *Uncertainty and Challenges in Contemporary Economic Behaviour*, Emerald Publishing, pp. 159-183, doi: 10.1108/978-1-80043-095-220201013.

Ying, R., Shou, Y. and Liu, C. (2021), "Prediction model of Dow Jones index based on LSTM-Adaboost", *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*, pp. 808-812.

Yue, H., Liao, H., Li, D. and Chen, L. (2021), "Enterprise financial risk management using information fusion technology and big data mining", *Wireless Communications and Mobile Computing, Hindawi*, Vol. 2021, pp. 1-13.

Zhang, X., Li, A. and Pan, R. (2016), "Stock trend prediction based on a new status box method and AdaBoost probabilistic support vector machine", *Applied Soft Computing*, Vol. 49, pp. 385-398.

Zhao, G., Cheng, L. and Hu, H. (2022), "Power marketing assistant decision-making method based on big data mining", *2022 International Conference on Big Data, Information and Computer Network (BDICN)*, *IEEE*, pp. 165-168.

Zhu, C., Kang, L. and Feng, W. (2021), "Predicting stock closing price with stock network public opinion based on AdaBoost-AAFSA-Elman model and CEEMDAN algorithm", *Journal of Shanghai Jiaotong University (Science)*, pp. 1-13, doi: 10.1007/s12204-021-2337-x.

Appendix

Table A1.
Predicted daily precision improvement of JKII prices in Indonesia's Islamic stock market using big data mining techniques

Days	JKII actual prices	Predicted JKII prices using AdaBoost	Predicted JKII prices using random forest	Predicted JKII prices using KNN	Predicted JKII prices using ANNs
March 03, 2022	590.17	590.124418	591.205625	592.484	588.9351547
April 01, 2022	594.16	594.2521658	594.223	595.57	592.4528272
April 04, 2022	600.34	600.6896415	601.4437833	602.788	602.6387391
April 05, 2022	605.86	605.3862121	604.7257381	607.436	603.1847489
April 06, 2022	604.57	604.6664198	604.7081508	602.766	607.6727891
April 07, 2022	604.24	604.389148	604.285269	606.602	607.4380509
April 08, 2022	611.26	611.3562178	612.3096262	609.598	614.1947918

Source: Predicted daily precision improvement of JKII prices in Indonesia's Islamic stock market using big data mining

Corresponding author
Mohammed Ayoub Ledhem can be contacted at: ledhem.edu@gmail.com