Title.

Stock Market Prediction using Artificial Intelligence: A Systematic Review of Systematic Reviews

Author names and affiliations.

Chinyang Lin (Coka) [a], João Alexandre Lobo Marques [a*]

Chinyang Lin (Coka)

[a] Faculty of Business and Law, University of Saint Joseph, Macau, China

E-mail:

supercoka@hotmail.com , lin.chin.yang@usj.edu.mo

João Alexandre Lobo Marques

[a] Faculty of Business and Law, University of Saint Joseph, Macau, China

E-mail:

alexandre.lobo@usj.edu.mo

[*]**Corresponding author.**

João Alexandre Lobo Marques

Present/permanent address:

Faculty of Business and Law, University of Saint Joseph,

Estrada Marginal da Ilha Verde, 14-17, Macau, China

E-mail:

alexandre.lobo@usj.edu.mo

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no competing interests.

Authors' contributions

The first author contributed by retrieving literature, extracting the data and. The second author contributed by conducting data analysis and verifying the data. All authors wrote the paper and read and approved the final manuscript.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Stock Market Prediction using Artificial Intelligence:

A Systematic Review of Systematic Reviews

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Abstract

There are many systematic reviews on predicting stock. However, each of them reveals a different portion of the hybrid AI analysis and stock prediction puzzle. The principal objective of this research was to systematically review and conclude the systematic reviews on AI and stock to provide particularly useful predictions for making future strategies for stock markets. Keywords that would fall under the broad headings of AI and stock prediction were looked up in two databases, Scopus and Web of Science. We screened 69 titles and read 43 systematic reviews which include more than 379 studies before retaining 10 of them.

Keywords: Machine Learning, Deep Learning, Support Vector Machines (SVM), Long Short-Term Memory (LSTM), Neural Networks (NN)

JEL Classification codes: C880, C890, G000, G100, G110, G170

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

1. Introduction

In recent years, researchers have shown an increased interest in predicting the stock price and beating the stock market with Artificial Intelligence (AI) **(Sun et al., 2016; Macchiarulo, 2018; Liang et al., 2020; Samitas et al., 2020; Mokhtari et al., 2021; Song & Jain, 2022; Chhajer et al., 2022)**. Different angles have been taken when discussing the prediction of AI technology and the stock market **(Petropoulos et al., 2022; Shmueli & Tafti, 2022)**. More hybrid technology and information can be associated with better prediction accuracy in stock markets **(Bustos & Pomares-Quimbaya, 2020; Kumar et al., 2022; Smyl, 2020)**. In contrast, the data sets used to train the algorithms of deep learning (DL) models are not large enough to give accurate predictions currently **(Bustos & Pomares-Quimbaya, 2020; Hewamalage et al., 2021)**. Many systematic reviews have synthesized statistical results concerning AI and stock price; however, taken separately, these reviews provide us with fragments of AI and the puzzle of stock prediction **(Makridakis et al., 2020)**. A recent literature review identified several gaps in the literature and emphasized the need to reconsider how we approach AI and stock using analysis from previous systematic reviews **(Pinto, Figueiredo, et al., 2021)**. We contend that a systematic review of systematic reviews (herein referred to as a meta-review) is required to have a general understanding of the most recent developments in AI and stock research. Throughout the search, methodologies specifically designed to comprehend the various approaches to AI and stock research will be used in this meta-review. It will thoroughly review, contrast, and conclude the strength of the evidence from previous systematic reviews. It will evaluate each review critically with the intention of influencing future AI and stock research design, strategy, and execution.

This study includes economics, statistics, finance, and computer science theories. Stock prices are driven by new information that cannot be obtained utilizing an analysis through the stock market, according to two significant financial theories: the Random Walk Model **(Fama et al., 1969; Fama, 1995)** and the Efficient Market Hypothesis **(Fama, 1965)**. Nevertheless, numerous researchers have disproved the underlying assumptions of these two hypotheses and demonstrated that the market could be partially forecast in accordance with socioeconomic theory and behavioral economics/finance **(Patel et al., 2015; Chong et al., 2017; Oliveira et al., 2017; Weng et al., 2017)**.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Recent research suggests that the methods of stock market analysis are split into mathematical and AI techniques. In this respect, mathematical technology refers to statistical tools, and AI technology refers to Machine Learning (ML) algorithms **(Januschowski et al., 2020)**. Furthermore, it has been observed that most of the selected studies utilize ML algorithms to analyze the performance of stock market prediction **(Kumar et al., 2022a)**. Previous research has established that ML, a subset of AI, includes Deep Learning (DL) **(Jakhar & Kaur, 2020)**. From this perspective, the various techniques utilized to predict stock prices can be preliminarily split into 1. traditional ML algorithms, and 2. DL and Neural Network (NN) **(Soni P., 2022)**. However, there are numerous methods among traditional ML, DL, and NN for predicting stock prices. One of the most significant current discussions in predicting the stock market is which methods are most frequently employed to forecast stock market prices **(Li & Bastos, 2020)**.

This study systematically reviews and concludes the systematic reviews for the stock market forecasting techniques, aiming to provide particularly useful predictions for making future stock market strategies. Therefore, the research questions (RQ) in this study focused on systematically reviewing the universality of AI techniques for stock market prediction, informational sources used, and performance metrics. While a variety of definitions of the term AI, ML, and DL have been suggested, this paper will use the definition suggested by **Jakhar & Kaur (2020)**. For the aspect of AI, incorporating human intelligence into machines is one way to define AI broadly, and any computer program with elements of human intelligence is referred to as AI. For the aspect of ML, a subset of AI called ML consists of all techniques that let computers learn from data without being explicitly programmed. The goal of ML is to train computers using the available data and methods. Machines learn how to make decisions using the data and information that have been processed. In essence, ML is just a method for making AI. For the aspect of DL, A subset of ML, DL includes the artificial neural network (ANN) that mimics the structure of biological neural networks seen in the brain. Every time the brain learns new information, it attempts to make sense of it by comparing it to previously learned knowledge. DL uses this strategy similar to how the brain organizes information by categorizing and labeling objects. Generally speaking, DL is more accurate than ML and performs exceptionally well on unstructured data, but it also needs a massive amount of training data and expensive hardware and software **(Jakhar & Kaur, 2020)**.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

2. Methods

To determine how researchers in the field conduct and report systematic literature reviews, we conducted a systematic review of systematic reviews that have been published in the public domain. This systematic review was written in accordance with the PRISMA 2020 Statement **(Page, McKenzie, et al., 2021)**, which consists of a checklist and a flow diagram. PRISMA Elaboration and Explanation **(Page, Moher, et al., 2021)** improves the usage, comprehension, and public awareness of the PRISMA 2020 Statement. The purpose and rationale for each checklist item are conveyed through examples and explanations.

This systematic review included all published articles of systematic reviews concerning stock and the application of AI technology. Kaplan & Haenlein (2019) define AI as "a system's ability to correctly interpret external data, to learn from such data, and to apply those learnings to achieve specific goals and tasks through flexible adaptation." We included reviews that approached stock from the stock market or price perspective.

A well-defined research methodology must be followed in addition to some specific criteria in order to choose the primary publications to be applied in this work. **Kitchenham & Charters (2007)** suggested three processes for creating a systematic review: planning, conducting, and results from the analysis.

2.1. PLANNING THE SYSTEMATIC REVIEW

As a result, the first stage must include a few inquiries that will need to be addressed after the systematic review, as well as the inclusion, exclusion, and quality criteria. This systematic review will answer the following questions in Table 1:

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

TABLE 1. Research Questions.

ID	Research Questions (RQ)
RQ1	Which AI methods and technologies are most commonly utilized to forecast stock market prices?
RQ2	Which informational sources are most frequently utilized to predict stock market prices?
RQ3	Which metrics are most popular employed to verify the performance of the predictive models?

The inclusion (IC) and exclusion (EC) criteria are presented in Table 2.

TABLE 2. Inclusion and Exclusion Criteria.

Inclusion Criteria (IC)	**Exclusion Criteria (EC)**
Searched in Title	Not Searched in Title
Published in the English Language	Not Published in the English Language
Studies of the US, UK, and Europe	Not Studies of the US, UK, and Europe
Published between 2009 and April 2022	Published before 2009
Works using AI as the primary technique	

2.2. CONDUCTING THE SYSTEMATIC REVIEW

The second stage involves retrieving relevant publications for the systematic review and choosing the works according to the predefined criteria.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

As a result, the Web of Science database was added to complement the prior platform since it is one of the oldest **(Paras et al., 2018)**, and the Scopus platform was employed to extract papers because it is a reference in academia **(Wang et al., 2016)**. The terms "*Systematic Review*," "*Systematic Literature Review*," and "*Stock*" were utilized as search descriptions. Possible variations were also adopted for this selection to include the most significant number of articles relevant to the themes. Thus, the following search term was utilized:

(("*Systematic Review*") **OR** ("*Systematic Literature Review*")) **AND** ("*Stock ?*")

Each search was configured to exclusively choose these terms on title documents for Web of Science and Scopus. Additionally, the term "articles" was utilized as a data collection limitation. The last search was run on April 11, 2022, and 69 articles were published on each platform using the keywords. The table which discriminates the number of studies mined through each database is presented in Table 3.

Table 3. A table that discriminates the number of studies mined through each database

Database	**Number of Studies**
Scopus	40
Web of Science	29

With the aid of Rayyan[1], a free web and mobile app that helps accelerate the initial abstract and title screening with the application of a semi-automated approach that incorporates excellent usability **(Ouzzani et al., 2016)**, duplicate publications could be eliminated, resulting in 43 articles.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Studies that were obviously not systematic reviews, such as empirical, descriptive, and conceptual papers, were removed after two independent reviewers separately evaluated the titles and abstracts of the records. The two reviewers then independently and thoroughly read the remaining full texts of papers to perform an eligibility evaluation. Disagreements among the reviewers were discussed and settled by consensus during this stage.

Following the reading of the abstract, the inclusion and exclusion criteria were applied, yielding 17 articles. Besides works belonging to the same authors and duplicated with the article of the book, the complete work was considered. In this way, one publication was excluded. Among the 16 articles, there are no studies in which we cannot retrieve the full text. The final step involved reading the remaining 16 articles while excluding any that did not meet the inclusion and exclusion standards. As a result, the following step will involve analyzing the final ten articles. Based on PRISMA Elaboration and Explanation **(Page, Moher, et al., 2021)**, the number of articles separated during the specified method is displayed in Figure 1.

[1]http://rayyan.qcri.org

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

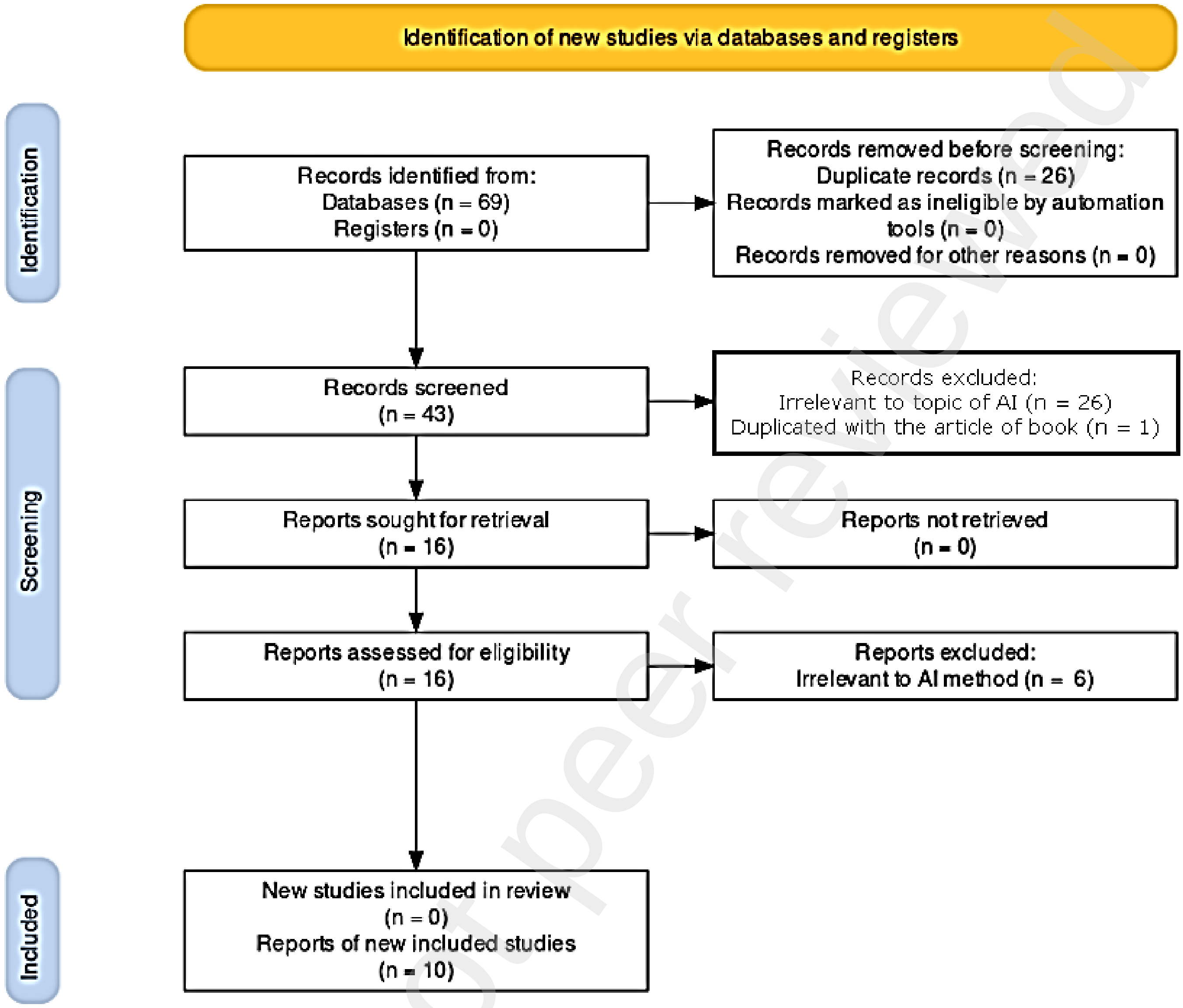

FIGURE 1. The number of articles separated during PRISMA.

One author extracted the data from the included studies, while a second author verified the data. Discussions between the two reviewers helped to settle disagreements. The data was then carefully extracted and coded from the included studies.

After thoroughly reviewing the articles chosen, Table 4 was filled with all pertinent data for each reference. Such as the aim of the review, the date range of included studies, type of review, search sources, included studies, and Journal/Conference adopted. Table 5 was filled with the retrieved attributes and AI

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

analysis/method/technology utilized for training or testing the forecast model, including algorithms, tools, metrics, and databases.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Table 4. Review characteristics of analyzed articles (All Pertinent Data for Each Reference)

Reference	Aim of Review	Date range of included studies	Type of review	Search sources	Included studies	Journal or Conference
Jabbar Alkubaisi G.A.A. (2017)	Define the relationship between stock market forecast accuracy and feature selection by addressing specific studies that utilize sentiment analysis on financial news and reports and research that utilizes sentiment analysis on Twitter using machine learning classifiers.	2011 to 2017	SR	No mentioned.	27	Journal of Theoretical and Applied Information Technology
Islam M.R. (2018)	Systematically review text mining techniques, methodologies, and principal component analysis that are utilized to help minimize dimensionality in the characteristics and outstanding features. Systematically review the most complex soft-computing approaches and techniques that are evaluated for their performance using electronic textual data in terms of analysis, comparison, and evaluation.	No mentioned.	SR	No mentioned.	No mentioned.	Indonesian Journal of Electrical Engineering and Computer Science
Bustos & Pomares-Quimbaya (2020)	Update systematic review of forecasting machine learning (ML) techniques utilized in the stock market, such as Deep Learning (DL), Text Mining Techniques, and Ensemble Techniques, including classification, characterization, and comparison.	2014 to 2018	SR	Scopus, Web of Science	53	Expert Systems with Applications
Ketsetsis A.P. (2020)	Systematically review primary studies that utilize DL techniques to predict stock markets in the European Union (EU).	2011 to 2019	SR	Google Scholar	12	2020 International Conference on Computational Science and Computational

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

						Intelligence (CSCI)
Li & Bastos (2020)	Collect, analyze, and review existing academic articles on financial time series forecasting using DL and technical analysis.	2017 to 2020	SR	Scopus, Web of Science, IEEE Xplore	34	IEEE Access
Nti et al. (2020)	Systematically and critically review the research works reported in academic journals using ML for stock market forecast. Systematically review studies on stock market predictions based on fundamental and technical analyst perspectives, resulting in a better understanding of the current state of the art and possible future directions.	2007 to 2018	SR	No mentioned.	122	Artificial Intelligence Review
Zavadzki et al. (2020)	Systematically review the literature on research involving stock market forecasting techniques and seek to identify research in which advanced computational models have been applied to the stock market, as well as to describe the main computational intelligence techniques utilized by such research.	2014 to 2018	SR	IEEE Xplore Digital Library, Science Direct, Scopus, Web of Science.	24	IEEE LATIN AMERICA TRANSACTIONS
Pinto et al. (2021)	Systematically review the selected papers using time series, text mining, and sentiment analysis applied to predict financial stock market behavior.	2015 to 2019	SR	ACM Digital Library, Google Scholar, IEEE Digital Library, Science @ Direct e Springer Link	57	2021 IEEE 24th International Conference on Computer-Supported Cooperative Work in Design (CSCWD)

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Kumar et al. (2022)	Systematically review stock market forecast methodologies and academic papers that suggest techniques, including calculating techniques, ML algorithms, performance factors, and top journals.	2002 to 2019	SR	IEEE, Springer, ScienceDirect, Scopus, MDPI	30	Materials Today: Proceedings
Soni P. (2022)	Investigate the various techniques utilized in stock price prediction, ranging from traditional ML and DL methods to neural networks and graph-based approaches.	2016 to 2021	SR	No mentioned.	20	Journal of Physics: Conference Series (JPCS)

Table 5. Most Utilized Review characteristics of analyzed articles (Stock Analyses and Predictor Techniques for Each Reference)

Reference	**Method of Stock Market Prediction**	**Informational Sources**	**Metrics**
Jabbar Alkubaisi G.A.A. (2017)	1. ML. 2. Sentiment Analysis. 3. Statistical Measurements.	1. Twitter 2. Timestamps (temporal feature) 3. Geographic location (Spatial feature)	No mentioned.
Islam M.R. (2018)	1. Principle Component Analysis (PCA). 2. Technical Approach in Text Mining, including Genetic Algorithms, DL, ML, Apriori-Like Algorithm, and Fuzzy Algorithm.	1. Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) 2. (1) Correlation among different stocks (2) Historical stock prices (3) Newspapers content	No mentioned.
Bustos & Pomares-Quimbaya (2020)	1. PCA. 2. Fundamental and Technical Analysis.	1. Structured data: (1)Market Information	1. Accuracy 2. Precision

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

	3. ML, including DL, Text Mining, and Ensemble Techniques. 4. Artificial Neural Network (ANN). 5. Support Vector Machine (SVM). 6. Bio-Inspired Computing. 7. Sentiment Analysis. 8. Social Network Analysis.	(2)Technical Indicators (3)Economic Indicators 2. Unstructured data: (1)News (2)Social Network (3)Blogs	3. Recall 4.F1-score
Ketsetsis A.P. (2020)	1. ML. 2. DL. 3. Time Series Analysis. 4. Text Mining. 5. GRU, GRU-SVM (Gated Recurrent Unit-SVM), SVM. 6. NN, CNN (Convolutional Neural Network), DNN (Deep neural network). 7. Long short-term memory (LSTM). 8. MLP (Multilayer Perceptron), Mixed ARMA-MLP (Auto Regression Moving Average-MLP). 9. Hybrid Fuzzy-Neural Network. 10. GA-SVR (Genetic Algorithm-Support Vector Regression).	1. Financial indicators 2. EUR/USD Exchange Rates 3. Gold Prices.	1. Mean Squared Error (MSE) 2. Accuracy 3. Mean Absolute Error (MAE).
Li & Bastos (2020)	1. Technical Analysis. 2. DL. 3. LSTM. 4. CNN. 5. Tools: Python, Tensorflow, NumPy, Pandas, Scikit-Learn, Keras, TA-Lib, and TA4J10.	Historical stock prices (most authors utilize daily data), e.g. 1. Yahoo Finance. 2. Wind. 3. Taiwan Stock Exchange.	1. Accuracy 2. Precision 3. Recall 4.F1-score
Nti et al. (2020)	1. Fundamental and Technical Analysis. 2. ML. 3. SVM. 4. ANN.	1. Historical stock prices and technical indicators in technical analysis	1. Mean Absolute Percentage Error (MAPE) 2. MSE 3. MAE

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

	5. Term Frequency-Inverse Document Frequency (TF-IDF). 6. Feature Selection Techniques. 7. Correlation Analysis. 8. Tool: MATLAB.	2. Financial ratios of the firm and unstructured nature of fundamental factors in fundamental analysis	4. Root Mean Squared Error (RMSE) 5. Correlation coefficient (R) 6. Volatility 7. Momentum 8. Accuracy 9. Precision 10. Recall 11. F-score 12. Normalized Mean Squared Error (NMSE) 13. Prediction of Change in Direction (POCID)
Zavadzki et al. (2020)	1. ANN. 2. Fuzzy. 3. Evolutionary Computation (Comp. Evol.). 4. Support Vector Regression (SVR). 5. K-Nearest Neighbor (KNN). 6. Principal Component Analysis (PCA). 7. Random Forest (RF). 8. Empirical Mode Decomposition (EMD). 9. SVM. 10. Adaptive Network-based Fuzzy Inference System (ANFIS)	1. Dow Jones indices 2. NASDAQ (National Association of Securities Dealers Automated Quotations) 3. TAIEX index (Taiwan Capitalization Weighted Stock Index)	1. RMSE 2. MAPE
Pinto et al. (2021)	1. Time Series Analysis (applicable method). 2. Text Mining (applicable method). 3. Sentiment Analysis (applicable method).	1. Microblogs, news, lexical dictionaries, and Twitter for text mining.	No mentioned.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

	4. ML and SVM (the most cited algorithm). 5. DL (the most commonly utilized technique). 6. NNs (the most commonly utilized techniques).	2. News and Twitter for sentiment analysis. 3. Time series	
Kumar et al. (2022)	1. ML algorithms, e.g., NN, ANN, SVM 2. Statistical Techniques, e.g., ARIMA (An Autoregressive Integrated Moving Average) and Clustering.	NASDAQ (most of the selected studies utilized dataset)	1. Accuracy 2. MSE 3. RMSE
Soni P. (2022)	1. Traditional ML Method 2. DL and NNs, e.g., Recurrent neural network (RNN) and LSTM. 3. Time-Series Analysis. 5. Graph-Based Analysis.	1. Historical stock prices 2. News 3. Technical indicators 4. Data from various online platforms such as Yahoo Finance and Twitter.	1. Accuracy 2. RMSE 3. MAPE 4. MAE

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

2.3. ANALYSIS OF RESULTS

The third and final process entails analyzing the chosen articles and classifying them according to predictor methodologies, informational sources, metrics employed to verify the effectiveness of the predictive model, and the limitations and future recommendations of the research.

2.3.1 Analysis based on predictor methodologies

This section tries to examine methods for predicting the trends or prices of the stock market. Once the focus of the work is statistical analyses and AI methods of prediction of the stock market, the analysis will cover Statistical Techniques, Time Series Analysis, Principal Component Analysis (PCA), Fundamental and Technical analysis, ML, DL, Text Mining, NN, ANN, SVM, Sentiment Analysis, LSTM technique, and others.

Because statistical techniques treat financial time series as linear systems **(Li & Bastos, 2020)**, some writers claimed that works based on statistical methods did not perform well and produced inferior results than models based on Artificial Intelligence (AI) **(Atsalakis & Valavanis 2009; Boyacioglu & Avci 2010; D. A. Kumar & Murugan 2013; Bisoi & Dash 2014)**.

A well-known method for condensing high-dimensional datasets and extracting the essential features from training inputs is principal component analysis (PCA). PCA reduced the quantity of data needed to train the models while simultaneously improving the accuracy of the predictions **(Bustos & Pomares-Quimbaya, 2020)**. Thus, **Islam M.R. (2018)** focused on alternative approaches and strategies for the usability of textual and numerical data on the stock price trend based on the PCA methodologies. **(Misra et al., 2018)** also

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

discovered that the accuracy of predictions made using the linear regression model increases when PCA is applied to choose the most pertinent components from the data.

The outcome of a systematic and critical review of 122 relevant research articles revealed that technical analysis accounted for 66% of the documents examined. In contrast, fundamental and combination analyses accounted for 23% and 11%, respectively **(Nti et al., 2020)**. Because it is more difficult to construct models that explain why a stock is moving, fundamental analysis is less frequently discussed in the literature **(Bustos & Pomares-Quimbaya, 2020)**.

In general, ML and DL produce better results than conventional time-series models because they take advantage of the modern computer's rising processing capability **(Siami-Namini et al., 2018)**. **Soni P. (2022)** examined the outcomes of several algorithms, such as sentiment analysis, time series analysis, and graph-based algorithms, in order to predict the stock values of various companies. In terms of accuracy, traditional ML algorithms continue to beat graph-based methods.

Most of the main techniques cited are ML algorithms **(Pinto, da Silva Figueiredo et al., 2021)**. The most commonly employed ML algorithms in stock market predictions are Decision Tree (DT), SVM, and ANN **(Nti et al., 2020)**. Furthermore, the most cited algorithm was SVM **(Pinto, da Silva Figueiredo, et al., 2021)**. Four ML methodologies, including Genetic Algorithms mixed with other techniques, ANN, SVM, and hybrid methods, were identified after **(Strader et al., 2020)** studied journal publications from the previous twenty years. The various techniques employed to predict stock prices can be broadly classified into four

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

groups. Typical ML Techniques incorporate established techniques like logistic and linear regression analysis. Recurrent neural network (RNN) and LSTM are frequently applied in NN and DL approaches **(Soni P., 2022)**.

Four ML ensemble techniques are trained on the artificial intelligence platform: random forest regression, reinforced regression tree, SVR ensemble, and neural network regression ensemble. The article on the forecasting models of stock market share employed various computational intelligence techniques, primarily ANN, Fuzzy, evolutionary computing algorithms, and Support Vector Regression **(Zavadzki et al., 2020)**. To achieve exact stock market predictions, the most extensively utilized techniques in 30 studies are ANN (24%) and NN (33%) **(Kumar et al., 2022b).**

DL application in stock market technical analysis is a relatively new concept **(Li & Bastos, 2020)**. DL is one of the most popular methods in ML. DL is also adequate for event-driven stock price movement prediction based on a mix of long-term and short-term occurrences. Traditional ML methods perform worse than DL approaches, while DL can still not investigate streaming data, distributed computing, and computing scalability **(Islam M.R., 2018)**.

In addition, the application of complex ML methods, such as ensemble models and DL, has grown in popularity. Ensemble models have demonstrated significant predictive power, outperforming other techniques such as SVM and ANN in some comparison studies. In general, DL models have not outperformed standard models. Likely, the data sets employed to train these algorithms are not large enough to give accurate predictions **(Bustos & Pomares-Quimbaya, 2020)**.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Nevertheless, the application of DL, NN, and SVM has distinguished itself and demonstrated a potential trend of producing accurate predictions among various ML methods **(Pinto, da Silva Figueiredo et al., 2021)**. DL techniques have led to their application in time-series forecasting, such as stock prices, by the research community **(Ketsetsis A.P., 2020)**. In this respect, RNN and LSTM are the best DL algorithms employed for the task. RNN has the advantage of capturing the context of the data as they are being trained **(Soni P., 2022)**. RNN is the most extensively studied by researchers, according to **Sezer et al. (2020)**, who surveyed DL approaches for forecasting financial time series.

Regarding the DL techniques, the scientific community revealed a definite preference for LSTM. The advantage of LSTM as a stock market forecasting approach in terms of frequency. This pattern is consistent with the global trend, demonstrating that the LSTM approach is the most useful for forecasting global stock markets **(Ketsetsis A.P., 2020)**. The LSTM network was utilized in most research because it can store memory and address the gradient vanishing problem, making it the perfect approach for time series forecasting **(Li & Bastos, 2020)**. Because they can correlate non-linear time series data in the delay state, LSTM performs reasonably well **(Soni P., 2022)**.

Furthermore, **Di Persio & Honchar (2016)** presented a combination of CNN and wavelets for projecting this S&P500 index by using closing prices as the testbed. They concluded that for predicting market indexes, CNN outperforms traditional neural networks. Moreover, social network analysis has proven effective for stock predictions using sentiment indexes and other derived series as inputs **(Bustos & Pomares-Quimbaya, 2020)**.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

The goal of fuzzy logic is to imitate human thought. Fuzzy logic enables the development of soft if-then rules in which the premises are represented by categories rather than precise values. A potent approach for learning rules from human experts is fuzzy logic. The most often utilized algorithm is the adaptive neuro-fuzzy inference system (ANFIS) **(Bustos & Pomares-Quimbaya, 2020)**.

One of the most common feature-representation techniques for textual data was the Term Frequency-Inverse Document Frequency (TF-IDF). However, feature selection techniques based on correlation analysis are utilized in 99% of the works assessed **(Nti et al., 2020)**.

According to the papers that disclose what tools were applied, most authors employed Python and Tensorflow to program the predictor based on past price data. Pandas, NumPy, Keras, Scikit-Learn, TA-Lib, and TA4J, were additional extensively applied tools **(Li & Bastos, 2020)**. However, another review indicated that MATLAB is the most popular modeling software for stock market forecasting **(Nti et al., 2020)**.

2.3.2 Analysis based on informational sources

For predicting share prices, only historical data was applied the past. Analysts now understand that numerous additional aspects are crucial in determining the stock price, making it inaccurate to rely solely on past data **(Soni P., 2022)**. Technical indicators are the most widely employed source of information for stock market forecasting. Technical indicators have been indicated to be the most accurate data. The input from social networks, on the other hand, helps the models perform better **(Bustos & Pomares-Quimbaya, 2020)**.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Time series is by far the most frequent feature set for stock price prediction models, with financial indicators coming in second **(Ketsetsis A.P., 2020)**.

Studies reveal that human sentiments and emotions can aid in predicting stock market returns in addition to historical financial information about companies or stock markets. Twitter is one of the primary information sources from social networks that are now accessible to everyone, and tweets from important people emotionally impact people, ultimately impacting their investing decisions **(Jabbar Alkubaisi G.A.A., 2017)**. In other words, the application of data from social media and websites is a combined source of information that allows for more accurate forecasting **(Pinto, Figueiredo, et al., 2021)**.

Furthermore, future research is likely to focus on identifying new sources of information that may be utilized in conjunction with technical analysis to forecast stock markets. Technical indicators, which have been demonstrated to be the most accurate data, are the most widely applied source of information for stock market forecasting. In other respects, the input from social networks helps the models perform better **(Bustos & Pomares-Quimbaya, 2020)**.

In contrast to technical indicators, fundamental ones are less frequently discussed in the literature because it is more difficult to construct models explaining why a stock is moving. The most frequently utilized data relates to macroeconomic time series, including, but not limited to, Gross Domestic Product (GDP), Customer Pricing Index (CPI), currency exchange rates, and interest rates **(Boyacioglu & Avci, 2010)**. Other information sources are just as familiar as financial news, but they are more challenging to utilize due to their unstructured character and erratic behavior. Techniques for text mining have been employed to deal with this

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

complexity. In addition, the news analysis is usually taken from three different sources: specialized media in finance, news in general, and news generated by the same company **(Bustos & Pomares-Quimbaya, 2020)**.

In another respect, the stock market returns can also be influenced by regional and temporal characteristics. The spatial characteristic may be other stock markets that have the potential to impact the local stock market, or it may be the various emotions of individuals from other geographic locations. Similar to this, a temporal effect depicts how something changes over time. People may have various perspectives at different times, and depending on their feelings at that particular moment, they may behave in different ways. Finally, all of these variables aid in our ability to forecast stock market returns **(Jabbar Alkubaisi G.A.A., 2017)**.

Although closing prices are the most frequently employed data, volume and ranges have also proven helpful in making predictions. Most research utilizes periods of 1000 days, which can be easily handled by most machine-learning algorithms **(Bustos & Pomares-Quimbaya, 2020)**.

Most authors choose Yahoo Finance as the informational source because of the simplicity with which they may obtain data using Yahoo Finance, a Python module. Most literary works employ daily data because it is simple and cost-free to collect this information from financial websites like Yahoo Finance **(Li & Bastos, 2020)**. Also, the NASDAQ dataset was employed in most of the selected research for stock market prediction and forecasting **(Kumar et al., 2022a)**.

Simple-Moving Average (SMA), Exponential Moving Average (EMA), Moving Average Convergence/Divergence rules (MACD), Relative-Strength Index (RSI), and Rate of Change (ROC) were indicated to be the most commonly utilized technical indicators for stock market prediction **(Nti et al., 2020)**.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Preprint not peer reviewed

The RSI, Commodity Channel Index (CCI), Williams R, and Stochastic Oscillators are the most popular oscillators **(Bustos & Pomares-Quimbaya, 2020)**.

The most popular trend indicators are Momentum (MOM) and Moving Averages. The performance of the previous day is summarized by the Simple Day Moving Average (SMA). When the crossover of trends occurs, it is employed in conjunction with various long-term averages for uptrend forecasting. Another well-known indicator that gives more weight to recent prices than historical prices is the Weighted Day Moving Average (WMA). Whether MOM is valued above or below zero, it recognizes trend lines. The MOM is determined by subtracting the two SMA **(Bustos & Pomares-Quimbaya, 2020)**.

2.3.3 Analysis based on the metrics employed to verify the performance of the predictive model

Since the review focused on the stock market, it is essential to examine the metrics employed to verify the accuracy of the predictive model. The research utilizes the metrics to evaluate their model and dataset **(Kumar et al., 2022a)**.

Ketsetsis A.P. (2020) indicated that Mean Squared Error (MSE) was the most commonly employed, appearing in 5 of the 12 papers, and was frequently accompanied by Accuracy and Mean Absolute Error (MAE). However**, Li & Bastos (2020)** and **Bustos & Pomares-Quimbaya (2020)** indicated that the most commonly employed measures to compare models are accuracy, precision, recall, and F1-score**.** Moreover, **Kumar et al. (2022a)** indicated that most researchers employ the accuracy performance metric to evaluate their model and dataset**.** On the other hand, **Nti et al. (2020)** indicated that Mean Square Error

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Preprint not peer reviewed

(MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) were the metrics that were most frequently employed in the literature. The fact that MSE and RMSE are preferable for assessing the performance of predictive models for short-term prediction explains their widespread application.

2.3.4 Analysis based on the limitations and future recommendations of research

The reviewed studies still have limitations when it comes to selecting and combining these data in a meaningful way **(Pinto, Figueiredo, et al., 2021)**. The study highlighted the limitations of the works and made several suggestions for future research.

Future research could concentrate on identifying new sources of data or information that can be utilized to supplement technical analysis to predict the movements of the stock market. More articles are expected to automatically find the best technical indicators **(Bustos & Pomares-Quimbaya, 2020)**. Researchers also can concentrate on merging sentiment analysis of stock-related data with the numeric value linked with previous stock value to forecast stock values **(Soni P., 2022)**. On the other hand, the possibility that the accuracy of ensemble algorithms may vary over different datasets from different continents presents another option for future research **(Ballings et al., 2015)**.

None of the one hundred and twenty-twos reviewed studies included social media sentiment, financial news, historical stock data, or macroeconomic data as input variables **(Nti et al., 2020)**. According to **Geva & Zahavi (2014)**, if all of these data sources are utilized as input for a predictive model, a better and higher level of prediction accuracy may be attained.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

In recent years, the application of time series, text mining, and sentiment analysis has increased. DL, SVM, and NN have all shone out. Despite advancements in research publications, the utilization of these three methodologies has some limits. The ability of the proposed models is limited by the usage of numerical data, textual information, and social media. Future research will look into combining time series, text mining, and sentiment analysis to automate better stock market prediction **(Pinto, Figueiredo, et al., 2021)**.

The main difficulty in stock market prediction is that most modern methods cannot be detected with the help of historical stock data. As a result, other factors, such as governmental policy choices and consumer attitudes, have an impact on stock markets **(Kumar et al., 2022a)**.

3. Discussion

The systematic review of systematic reviews is a comprehensive and systematic research study that aims to identify, evaluate, and synthesize the findings of all existing systematic reviews on a particular research topic. The research contribution of such a study could include the following:

1. Providing a thorough overview of the current state of knowledge on the research topic: A systematic review of systematic reviews can give a thorough overview of the current state of knowledge on the research topic by gathering and synthesizing the findings of all relevant systematic reviews.
2. Identifying gaps in the research: A systematic review of systematic reviews can identify gaps in the research topic, highlighting areas where further research is needed.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

3. Identifying areas of agreement and disagreement among the existing systematic reviews: A systematic review of systematic reviews can identify areas of agreement and disagreement among the existing systematic reviews, assisting in the clarification of conflicting results and pointing out areas where additional research is required to settle disagreements.
4. Guiding future research: A systematic review of systematic reviews can guide future research on the research topic by identifying areas where further research is needed and suggesting potential directions for future study.
5. Improving the quality of future research: By providing a comprehensive overview of the current state of knowledge on the research topic and identifying gaps in the research, a systematic review of systematic reviews can help improve the quality of future research by ensuring that it is well-informed and addresses important questions that have not yet been adequately addressed.

The statistical results relating to AI and stock price have been summarized in previous systematic reviews. However, these literature reviews provide us with fragments of AI and the puzzle of stock prediction. One of the aims of this study was to thoroughly analyze and wrap up the systematic reviews on AI and stocks to make forecasts that would be especially helpful when developing future stock market strategies. The research questions in this study sought to determine the most common AI methods, informational sources, and performance metrics for stock market prediction.

Based on the analyses performed in previous subsections and the data obtained in Table 5, it is possible to conclude the most commonly utilized methods for this review, the study trend over time, and several potential gaps to be explored in future works.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Although this systematic review limited publication years between 2009 and April 2022, the remaining articles dated from 2017 to 2022 after applying the inclusion and exclusion criteria, indicating that research conducted by AI with prediction analysis for the stock market is comparatively modern. One unanticipated result was that most articles about stock are relevant studies of AI methods, although the description of AI technology is not included in the search term. As illustrated in Figure 2, the number of publications has increased over time.

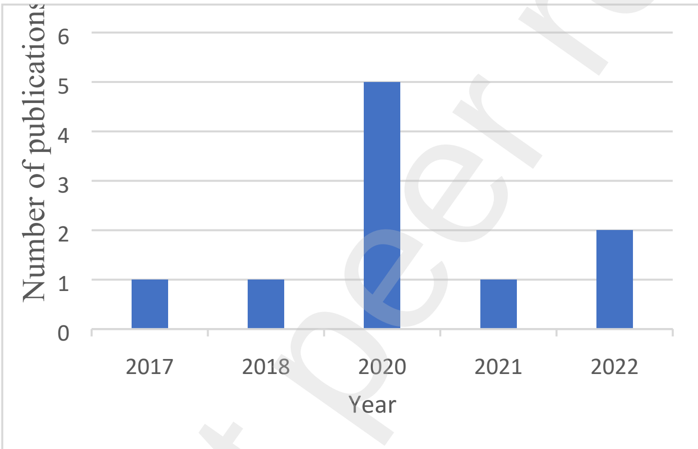

FIGURE 2. The number of publications per year.

Another unexpected and significant point to consider is the AI technique studied in each proposed systematic review, which answers the first question **(RQ1 - Which AI methods and technologies are most commonly utilized to forecast stock market prices?)**. In conclusion of the analysis of results, SVM, LSTM, and NN, including ANN, CNN, and RNN, are the most popular AI predictor approaches among different reviews. This study supports evidence from previous observations. SVM is the most frequently cited algorithm **(Pinto, da Silva Figueiredo, et al., 2021)**, and the scientific community preferred LSTM **(Ketsetsis A.P., 2020)**. The most often used techniques are the first, NN, and the second, ANN. **(Kumar et al., 2022b)**. RNN is the most extensively explored by researchers, according to **Sezer et al. (2020)**. CNN outperforms traditional

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

neural networks in predicting market indexes **(Di Persio & Honchar, 2016)**. This inconsistency may be due to the differences in data sets **(Bustos & Pomares-Quimbaya, 2020)** and the performance of learning methods **(Islam M.R., 2018)**. In this respect, one point that must be clarified is that DL is a subset of ML, which is a subset of AI **(Jakhar & Kaur, 2020)**. In general, SVM belongs to traditional Machine Learning **(Cortes & Vapnik, 1995)**; on the other hand, NN and LSTM, which is a kind of RNN **(Hochreiter & Schmidhuber, 1997)**, belong to DL. Another interesting point about which tools were utilized to create the forecaster based on stock price is that MATLAB is the most popular software besides Python and Tensorflow. These findings can help us to understand the relationship between data sets and prediction methods. In future investigations, it may be able to use diverse informational sources with different prediction approaches to maximize performance.

Responding to the second question (**RQ2 - Which informational sources are most frequently utilized to predict stock market prices?**), the most common feature set for stock price prediction models is time series of historical stock prices, followed by financial and technical indicators. Furthermore, closing prices are the most frequently employed data of historical stock prices, and SMA, EMA, MACD, RSI, and ROC are the most commonly utilized technical indicators. In addition, it is worth mentioning that social media sentiment, financial news, and macroeconomic data are also worthy informational sources for predicting stock prices. The present results are significant in at least two major respects. The stock market returns can also be influenced by regional and temporal characteristics **(Jabbar Alkubaisi G.A.A., 2017)**; on the other hand, the use of data sources is influenced by their complexity and difficulty. Therefore, a further study with more comprehensive thinking on the above variables is suggested.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Finally, answering the last question (**RQ3 - Which metrics are most popular employed to verify the performance of the predictive models?**), accuracy is generally the most popular and common metric employed to compare models. On the other hand, MSE is also a standard metric to verify the performance of the models. These findings suggest that accuracy and MSE are widely and easily applied. Therefore, accuracy and MSE can be used as the preliminary metrics of prediction performance. Further studies, which take more metrics such as MAE, RMSE, and MAPE into account, will need to be undertaken.

The most often mentioned articles regarding the gaps and suggested future work is related to identifying new sources of data or information **(Bustos & Pomares-Quimbaya, 2020)** and combining time series, text mining, and sentiment analysis to predict better the stock market **(Kumar et al., 2022a)**. In other words, the more information and hybrid technologies, the better performance of the predictive models.

4. Conclusion

Systematic reviews of systematic reviews are a type of literature review that aims to synthesize the findings of multiple systematic reviews on a particular topic. These reviews are increasingly being used to provide a comprehensive and up-to-date overview of the evidence on a particular topic and identify research gaps.

There are several areas where further research on systematic reviews of systematic reviews could be directed:

1. Methodology: There is an ongoing debate about the best methods for conducting systematic reviews of systematic reviews, including how to identify and select relevant reviews, how to assess the quality of individual reviews, and how to synthesize the findings. Researchers could focus on developing and evaluating new methods for conducting these types of reviews.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Preprint not peer reviewed

2. Data extraction and synthesis: There is a need for further research on extracting and synthesizing data from multiple systematic reviews, particularly in cases where the reviews use different methods or report on different outcomes. Researchers could investigate methods for synthesizing data from reviews that use different study designs or that report on different populations.

3. Reporting standards: There is a need for more consistent reporting standards for systematic reviews of systematic reviews, including guidelines for reporting on the methods used, the results, and the implications of the review. Researchers could work on developing and evaluating reporting standards for these types of reviews.

4. Impact: There is a need for research on the impact of a systematic review of systematic reviews on practice and policy. Researchers could investigate how these reviews inform decision-making and what factors influence their uptake.

5. Integration with other research synthesis methods: There is potential for systematic reviews of systematic reviews to be integrated with other research synthesis methods, such as network meta-analysis or realist synthesis. Researchers could investigate the benefits and challenges of integrating these approaches.

The main goal of the current study was to review the systematic literature reviews on the Prediction of Stock Markets using AI technologies, including ML and DL methods. A total of 10 papers were selected based on the research approach proposed for this systematic review. As a result, four essential points of view served as the foundation for analysis and discussion: predictor methodologies, informational sources, performance metrics of the predictive model, limitations, and future recommendations. This study has found that generally, SVM, LSTM, and NN, including ANN, CNN, and RNN are the most utilized predictor methodologies, the time series of historical closing stock prices are the most employed informational sources, the most employed performance metrics of the predictive model is accuracy, and the future recommendations for the prediction of a stock price is to utilize and combine as more information and technologies as possible. This new understanding should help to improve the accuracy of predictions of the stock market.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

The major limitation of this study is that the systematic review includes the articles searched in the title for a briefer review. Notwithstanding the relatively limited articles, this work offers brief and valuable insights into stock price predictions using AI technologies. Further research might explore more articles searched in more extensive fields, including searching the terms of keywords in the title, abstract, and keywords plus. Another critical and practical implication is that most research of stock prediction studies on the index of study. Thus, further research should be undertaken to forecast individual company stock.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Abbreviations

Abbreviations/Acronym	Full name
AI	Artificial Intelligence
ANFIS	Adaptive Network-based Fuzzy Inference System
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
ARMA	Auto Regression Moving Average
CCI	Commodity Channel Index
CNN	Convolutional Neural Network
Comp. Evol.	Evolutionary Computation
CPI	Customer Pricing Index
DL	Deep Learning
DNN	Deep Neural Network
DT	Decision Tree
EC	Exclusion Criteria
EMA	Exponential Moving Average
EMD	Empirical Mode Decomposition
GA-SVR	Genetic Algorithm-Support Vector Regression
GDP	Gross Domestic Product
GRU	Gated Recurrent Unit
IC	Inclusion Criteria
KNN	K-Nearest Neighbor
LSTM	Long short-term memory
MACD	Moving Average Convergence/Divergence rules
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLP	Multilayer Perceptron
MOM	Momentum
MSE	Mean Squared Error
NASDAQ	National Association of Securities Dealers Automated Quotations
NMSE	Normalized Mean Squared Error
NN	Neural Network
PCA	Principle Component Analysis
POCID	Prediction of Change in Direction
RF	Random Forest
RMSE	Root Mean Squared Error
ROC	Rate of Change
RQ	Research Question
RSI	Relative-Strength Index

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

SMA	Simple-Moving Average
SVM	Support Vector Machines
SVR	Support Vector Regression
TAIEX index	Taiwan Stock Exchange Capitalization Weighted Stock Index
TF-IDF	Term Frequency-Inverse Document Frequency
WMA	Weighted Day Moving Average

Declarations

Availability of data and materials

Data used in this paper were collected from Scopus and Web of Science Core Collection.

Competing interests

The authors declare that they have no competing interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

The first author contributed by retrieving literature, extracting the data and. The second author contributed by conducting data analysis and verifying the data. All authors wrote the paper and read and approved the final manuscript.

Acknowledgments

Not applicable.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

References

Ballings, M., Van den Poel, D., Hespeels, N., & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications*, *42*(20), 7046–7056. https://doi.org/10.1016/j.eswa.2015.05.013

Boyacioglu, M. A., & Avci, D. (2010). An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of the Istanbul stock exchange. *Expert Systems with Applications*, *37*(12), 7908–7912. https://doi.org/10.1016/j.eswa.2010.04.045

Bustos, O., & Pomares-Quimbaya, A. (2020). Stock market movement forecast: A Systematic review. *EXPERT SYSTEMS WITH APPLICATIONS*, *156*.

Chhajer, P., Shah, M., & Kshirsagar, A. (2022). The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction. *Decision Analytics Journal*, *2*, 100015. https://doi.org/10.1016/j.dajour.2021.100015

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, *83*, 187–205. https://doi.org/10.1016/j.eswa.2017.04.030

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

Di Persio, L., & Honchar, O. (2016). Artificial neural networks architectures for stock price prediction: Comparisons and applications. *International Journal of Circuits, Systems and Signal Processing*, *10*(2016), 403–413.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*(1), 34–105. https://doi.org/10.1086/294743

Fama, E. F. (1995). Random walks in stock market prices. *Financial Analysts Journal*, *51*(1), 75–80. https://doi.org/10.2469/faj.v51.n1.1861

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, *10*(1), 1–21. https://doi.org/10.2307/2525569

Geva, T., & Zahavi, J. (2014). Empirical evaluation of an automated intraday stock recommendation system incorporating both market data and textual news. *Decision Support Systems*, *57*, 212–223. https://doi.org/10.1016/j.dss.2013.09.013

Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent Neural Networks for Time Series Forecasting: Current status and future directions. *International Journal of Forecasting*, *37*(1), 388–427. https://doi.org/10.1016/j.ijforecast.2020.06.008

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Islam M.R., V. V., Al-Shaikhli I. F. ,. Nor R. B. M. (2018). *Technical approach in text mining for stock market prediction: A systematic review*. *10*(2). https://doi.org/10.11591/ijeecs.v10.i2.pp770-777

Jabbar Alkubaisi G.A.A., H. H., Kamaruddin S. S. (2017). *A systematic review on the relationship between stock market prediction model using sentiment analysis on twitter based on machine learning method and features selection*. *95*(24).

Jakhar, D., & Kaur, I. (2020). Artificial intelligence, machine learning and deep learning: Definitions and differences. *Clinical and Experimental Dermatology*, *45*(1), 131–132. https://doi.org/10.1111/ced.14029

Januschowski, T., Gasthaus, J., Wang, Y., Salinas, D., Flunkert, V., Bohlke-Schneider, M., & Callot, L. (2020). Criteria for classifying forecasting methods. *International Journal of Forecasting*, *36*(1), 167–177. https://doi.org/10.1016/j.ijforecast.2019.05.008

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Ketsetsis A.P., T. D., Kourounis C. ,. Spanos G. ,. Giannoutakis K. M. ,. Pavlidis P. ,. Vazakidis D. ,.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Champeris T. ,. Thomas D. (2020). *Deep Learning Techniques for Stock Market Prediction in the European Union: A Systematic Review*.

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*.

Kumar, D., Sarangi, P. K., & Verma, R. (2022a). A systematic review of stock market prediction using machine learning and statistical techniques. *MATERIALS TODAY-PROCEEDINGS*, *49*(8), 3187–3191. https://doi.org/10.1016/j.matpr.2020.11.399

Kumar, D., Sarangi, P. K., & Verma, R. (2022b). A systematic review of stock market prediction using machine learning and statistical techniques. *MATERIALS TODAY-PROCEEDINGS*, *49*(8), 3187–3191. https://doi.org/10.1016/j.matpr.2020.11.399

Li, A. W., & Bastos, G. S. (2020). Stock Market Forecasting Using Deep Learning and Technical Analysis: A Systematic Review. *IEEE ACCESS*, *8*, 185232–185242. https://doi.org/10.1109/ACCESS.2020.3030226

Liang, C., Tang, L., Li, Y., & Wei, Y. (2020). Which sentiment index is more informative to forecast stock market volatility? Evidence from China. *International Review of Financial Analysis*, *71*, 101552. https://doi.org/10.1016/j.irfa.2020.101552

Macchiarulo, A. (2018). PREDICTING AND BEATING THE STOCK MARKET WITH MACHINE LEARNING AND TECHNICAL ANALYSIS. *Journal of Internet Banking and Commerce*, *23*(1), 1–22.

Makridakis, S., Hyndman, R. J., & Petropoulos, F. (2020). Forecasting in social settings: The state of the art. *International Journal of Forecasting*, *36*(1), 15–28. https://doi.org/10.1016/j.ijforecast.2019.05.011

Misra, M., Yadav, A. P., & Kaur, H. (2018). Stock market prediction using machine learning algorithms: A classification study. *2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE)*, 2475–2478. https://doi.org/10.1109/ICRIEECE44171.2018.9009178

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Mokhtari, S., Yen, K. K., & Liu, J. (2021). Effectiveness of Artificial Intelligence in Stock Market Prediction based on Machine Learning. *International Journal of Computer Applications*, *183*(7), 1–8. https://doi.org/10.5120/ijca2021921347

Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). A systematic review of fundamental and technical analysis of stock market predictions. *ARTIFICIAL INTELLIGENCE REVIEW*, *53*(4), 3007–3057. https://doi.org/10.1007/s10462-019-09754-z

Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications*, *73*, 125–144. https://doi.org/10.1016/j.eswa.2016.12.036

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, n160. https://doi.org/10.1136/bmj.n160

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, *42*(1), 259–268. https://doi.org/10.1016/j.eswa.2014.07.040

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S.,

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., … Ziel, F. (2022). Forecasting: Theory and practice. *International Journal of Forecasting*, *38*(3), 705–871. https://doi.org/10.1016/j.ijforecast.2021.11.001

Pinto, N., da Silva Figueiredo, L., & Garcia, A. C. (2021). Automatic Prediction of Stock Market Behavior Based on Time Series, Text Mining and Sentiment Analysis: A Systematic Review. *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 1203–1208. https://doi.org/10.1109/CSCWD49262.2021.9437732

Pinto, N., Figueiredo, L. da S., & Garcia, A. C. (2021). *Automatic Prediction of Stock Market Behavior Based on Time Series, Text Mining and Sentiment Analysis: A Systematic Review—PROCEEDINGS OF THE 2021 IEEE 24TH INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED COOPERATIVE WORK IN DESIGN (CSCWD)*. 1203–1208.

Samitas, A., Kampouris, E., & Kenourgios, D. (2020). Machine learning as an early warning system to predict financial crisis. *International Review of Financial Analysis*, *71*, 101507. https://doi.org/10.1016/j.irfa.2020.101507

Shmueli, G., & Tafti, A. (2022). How to "improve" prediction using behavior modification. *International Journal of Forecasting*, S0169207022001066. https://doi.org/10.1016/j.ijforecast.2022.07.008

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1394–1401. https://doi.org/10.1109/ICMLA.2018.00227

Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, *36*(1), 75–85. https://doi.org/10.1016/j.ijforecast.2019.03.017

Song, G. H., & Jain, A. (2022). Can artificial intelligence beat the stock market? *Studies in Economics and Finance*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/SEF-03-2022-0133

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351

Soni P., K. D., Tewari Y. (2022). *Machine Learning Approaches in Stock Price Prediction: A Systematic Review*. *2161*(1). https://doi.org/10.1088/1742-6596/2161/1/012065

Strader, T. J., Rozycki, J. J., Root, T. H., & Huang, Y.-H. J. (2020). Machine learning stock market prediction studies: Review and research directions. *Journal of International Technology and Information Management*, *28*(4), 63–83.

Sun, A., Lachanski, M., & Fabozzi, F. J. (2016). Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction. *International Review of Financial Analysis*, *48*, 272–281. https://doi.org/10.1016/j.irfa.2016.10.009

Weng, B., Ahmed, M. A., & Megahed, F. M. (2017). Stock market one-day ahead movement prediction using disparate data sources. *Expert Systems with Applications*, *79*, 153–163. Scopus. https://doi.org/10.1016/j.eswa.2017.02.041

Zavadzki, S., Kleina, M., Drozda, F., & Marques, M. (2020). Computational Intelligence Techniques Used for Stock Market Prediction: A Systematic Review. *IEEE LATIN AMERICA TRANSACTIONS*, *18*(4), 744–755. https://doi.org/10.1109/TLA.2020.9082218

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4341351