https://acadpub.com/ics

Insights in
Computer Science

Artificial Intelligence Techniques for Stock and Market Index Prediction: A PRISMA-Based Systematic Review (2018-2024)

Mohamed B. Traoré, Kawtar Tikito, Saliha Assoul

Systems Engineering and Digital Transformation Laboratory, National School of Mines of Rabat, Rabat 10120, Morocco

About the Article

Open Access

Systematic Review

How to Cite:
M. B. Traoré, K. Tikito and S. Assoul, "Artificial intelligence techniques for stock and market index prediction: A PRISMA-based systematic review (2018-2024)," *Insights Comput. Sci.*, vol. 2, pp. 1-12. 2026.

Keywords:
Artificial intelligence, data driven approach, financial prediction, forecasting models, PRISMA

Corresponding author:
Mohamed B. Traoré,
Systems Engineering and Digital Transformation Laboratory, National School of Mines of Rabat, Rabat 10120, Morocco

ABSTRACT

Background and Objective: Financial asset prices are shaped by numerous endogenous and exogenous factors, including macroeconomic conditions, political events and market sentiment. With advances in artificial intelligence and the availability of large-scale data, novel techniques have emerged for financial forecasting. This study aimed to systematically review recent methodologies for predicting stock and market index prices.

Materials and Methods: A Systematic Literature Review (SLR) was conducted to identify, screen and analyze relevant scientific publications. Searches were performed in the Scopus and IEEE Xplore databases, yielding 32 eligible studies published primarily since 2018, comprising five review articles and 27 original research papers.

Results: The analysis reveals that hybrid modeling approaches are the most frequently adopted techniques in recent studies. Forecasting is predominantly focused on short-term horizons and a diverse set of input features-including technical, fundamental and sentiment-based indicators-is employed across models.

Conclusion: Recent research trends emphasize hybrid and short-term forecasting models for financial assets, reflecting the growing integration of artificial intelligence with financial data analytics. These findings provide a structured overview of current practices and highlight directions for future research in financial market prediction.

INTRODUCTION

Accurate prediction of stock and index prices is a critical task in financial markets, as it directly supports risk management, investment decision-making and algorithmic trading strategies. The ability to forecast future price movements enables investors to optimize portfolio allocation, hedge risks and anticipate market trends.

A stock represents a share of ownership in a company and entitles its holder to a claim on the firm's capital, potentially including voting rights and dividend payments depending on the stock type. In contrast, indices measure the performance of a group of companies, a specific sector, or an entire market. Financial indices serve multiple functions: They track market performance (e.g., the S&P 500 or CAC 40), provide benchmarks for evaluating investment funds and act as underlying assets for financial instruments such as Exchange-traded Funds (ETFs) and derivatives.

Investors are primarily concerned with the future evolution of asset prices and therefore seek reliable forecasting methods to manage risk effectively. A wide range of techniques has been proposed in the literature for predicting stock and index prices. Traditionally, statistical models such as the Multi-Layer Coupled Hidden Markov Model (MCHMM)[1] and ARFIMA[2] have been widely applied. With the expansion of the Internet and the advent of the Big Data era, new predictive approaches have gained prominence. Machine learning methods have been explored in previous studies; for example, Nam and Seong[3] employed Multiple Kernel Learning (MKL), while Zhang et al.[4] utilized coupled matrix and tensor factorization. Deep learning techniques have also become increasingly prevalent, including the LSTM model reported by Shah et al.[5] and the Associate Network proposed by Ding and Qin[6]. In addition, ensemble and decision-fusion methods have been applied to financial forecasting, as demonstrated by Jiang et al.[7] and Weng et al.[8]. Hybrid approaches that integrate multiple techniques are also gaining attention; For instance, Souma et al.[9] combined NLP, RNN and LSTM, whereas Huang et al.[10] integrated LSTM, BERT and GNN.

Given the growing volume of research on stock and index price prediction, a comprehensive literature review is essential to identify research trends and the most frequently employed techniques. However, most existing reviews focus exclusively on stocks. For example, Soni et al.[11] presented a semantic review of stock prediction methods, grouping them into four categories: traditional machine learning, deep learning, neural networks and graph-based approaches and discusses the associated challenges. Similarly, Mintarya et al.[12] reviewed 30 studies on financial prediction and reported that neural networks are the most commonly applied models. Kumar et al.[13] analyzed 30 research papers on stock price prediction and classified techniques according to different machine learning models. Finally, Lin and Marques[14] conducted a systematic review of review papers, examining 10 systematic reviews from Scopus and Web of Science and identified SVM, LSTM and ANN as the most widely used models for stock market prediction.

Another limitation of existing reviews is their focus on a restricted set of study characteristics. For instance, Soni et al.[11] examined only the algorithms and evaluation metrics used, while Mintarya et al.[12] provided an overview of machine learning techniques and trends in stock prediction. Kumar et al. [13] considered calculation methods, algorithms, performance metrics and publication venues and Lin and Marques[14] extended the analysis to include prediction methods, model input sources and evaluation metrics.

In practice, additional characteristics are of interest, including the temporal window of historical data, the type of prediction (trend or future price) and the prediction horizon. Accordingly, the objective of this study was to provide a systematic review of techniques used since 2018 for predicting financial instruments, with a specific focus on stocks and indices. To achieve this, PRISMA methodology was adopted for systematic literature reviews.

Contributions of the paper: The main contributions of this study are as follows:

- A systematic literature review of techniques used for stock and index prediction from 2018 to 2024
- A comprehensive analysis of key characteristics of the reviewed studies, including publication trends, types of prediction techniques and models, evaluation metrics, input data types, historical time windows, data sources, underlying stocks and indices and the nature and horizon of the predictions

RESEARCH METHODOLOGY

This section outlines the stages of the methodology adopted for conducting the literature review. It begins with the formulation of research questions, followed by the identification of keywords and construction of search queries to retrieve relevant studies from scientific databases. Subsequently, inclusion and exclusion criteria are defined. Finally, a synthesis of the selected articles is presented.

Formulation of the research questions: The first step involves defining the research questions guiding this study, which are as follows:

- What techniques are most widely used in recent years for predicting stock and index prices and which models are most frequently applied?
- What types of input data are considered by these techniques and what historical time windows are employed?
- What types of predictions are performed and what are the corresponding prediction horizons?
- Which evaluation metrics are used to assess model performance?

Definition of keywords and search query: In this stage, relevant keywords were identified and used to formulate a search query aimed at efficiently retrieving pertinent articles from selected scientific databases.

Given that the objective of this study is to identify recent techniques used in financial asset prediction, the analysis was restricted to two types of financial instruments: stocks and indices.

The final search query was formulated as follows:

- (Stock OR Index) AND (Pricing OR Forecasting OR Prediction) AND (Method OR Approach OR Technique)

Selection of scientific databases: Two major scientific databases were selected to ensure comprehensive coverage of the research topic: Scopus and IEEE Xplore.

Article search based on keywords: The formulated search query was applied to the Title, Keywords and Abstract fields in both databases. This process yielded the following results:

- 2,821 documents from Scopus
- 241 documents from IEEE Xplore

Definition of inclusion and exclusion criteria

Year of publication: Only documents published from 2018 onward were considered. After applying this criterion, the number of documents was reduced to:

- 1,864 documents from Scopus
- 81 documents from IEEE Xplore

Type of document and number of citations: Two additional filtering steps were applied. First, only research articles and review papers were retained. Second, a minimum citation threshold was imposed to select influential publications. After applying these criteria, the retained articles were:

- 49 articles from Scopus with at least 75 citations
- 70 articles from IEEE Xplore with at least 3 citations

The higher citation threshold for Scopus reflects the substantially larger volume of publications indexed in this database.

Relevance to domain and removal of duplicates: Only articles directly related to the finance domain and relevant to the research questions were retained. Subsequently, duplicate records across both databases were removed. After this step, the remaining articles were:

- 18 articles from Scopus
- 14 articles from IEEE Xplore

Table 1 summarizes the different stages of the inclusion and exclusion process. In total, 32 articles were retained for final analysis.

Table 2 presents the 32 selected articles, comprising 5 review papers and 27 research articles. For each article, the table reports its reference number in this study, title, year of publication and a link to the corresponding source. Table 2 presents 32 selected articles, comprising 5 journals and 27 research articles.

Full-text reading and synthesis: In the final stage, all 32 selected articles were thoroughly reviewed and synthesized. The outcomes of this synthesis are summarized in Tables 3, 5 and 6. For each study, key information was extracted, including the employed techniques and model types, evaluation metrics, types of input features used by machine learning models, historical data windows and sources, the underlying stocks or indices and the type and horizon of the predictions.

RESULTS AND DISCUSSION

Analysis of articles by publication year: Figure 1 illustrates the distribution of the 32 articles included in this study by publication year. The number of publications varies across years. All selected articles were published from 2018 onward, with the highest number appearing in 2020 (9 articles), followed by 2019 (8 articles), 2023 (4 articles) and 2018, 2021 and 2022 (3 articles each). The lowest number was observed in 2024, with 2 articles.

The peak in publications in 2020 may be attributed to the COVID-19 crisis, which had a substantial impact on financial markets and stimulated increased research activity in this area.

Analysis of articles by type: Figure 2 presents the distribution of article types. Among the 32 articles, two main categories are identified: Literature reviews and research articles. Research articles constitute the majority, with 27 papers (approximately 85%), whereas the remaining 5 papers (15%) are review articles. This imbalance is expected, as review papers typically synthesize findings from multiple primary studies.

Analysis of review articles: This subsection examines the five review articles included in the study, which focus on financial market forecasting.

Hu et al.[18] reviewed deep learning techniques for stock and foreign exchange prediction based on 86 papers published since 2015 and indexed in the Digital Bibliography & Library Project (DBLP). They reported that recent approaches frequently combine LSTM with other models.

Table 1: Number of selected articles according to inclusion and exclusion criteria

Inclusion / Exclusion criteria	Comment	Result
Application of search query	Applied to Title, Keywords and Abstract fields in each database	2821 papers from Scopus and 241 papers from IEEE Xplore
Year of publication	Only papers published between 2018 and 2024 were included	1864 papers from Scopus and 81 papers from IEEE Xplore
Type of document and number of citations	Only review and research articles were retained. A minimum of 75 citations for Scopus and 3 citations for IEEE Xplore was required	49 papers from Scopus and 70 papers from IEEE Xplore
Paper related domain and removal of duplicates	Only papers related to the finance field were retained	18 papers from Scopus and 14 papers from IEEE Xplore

Table 2: List of articles selected for the study (n = 32), including 5 review papers and 27 research articles

References	Title	Year of publication	Link
[15]	A novel graph convolutional feature based convolutional neural network for stock trend prediction	2020	#1
[9]	Enhanced news sentiment analysis using deep learning	2019	#2
[16]	Stock Market Trend Prediction Using High-Order Information of Time Series	2019	#3
[17]	An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market	2020	#4
[3]	Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market	2019	#5
[18]	A Survey of Forex and Stock Price Prediction Using Deep Learning	2021	#6
[8]	Predicting short-term stock prices using ensemble methods and online data sources	2018	#7
[19]	EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction	2018	#8
[20]	Stock price prediction using deep learning and frequency decomposition	2021	#9
[21]	Stock Closing Price Prediction using Machine Learning Techniques	2020	#10
[7]	An improved Stacking framework for stock index prediction by leveraging tree-based ensemble models and deep learning algorithms	2020	#11
[2]	Fractional Neuro-Sequential ARFIMA-LSTM for Financial Market Forecasting	2020	#12
[22]	Sentiment analysis on stock social media for stock price movement prediction	2019	#13
[4]	Improving stock market prediction via heterogeneous information fusion	2018	#14
[23]	Stock Market prediction on High frequency data using Long-Short Term Memory	2020	#15
[6]	Study on the prediction of stock price based on the associated network model of LSTM	2019	#16
[5]	Stock Market Analysis: A Review and Taxonomy of Prediction Techniques	2019	#17
[24]	Systematic analysis and review of stock market prediction techniques	2019	#18
[25]	A Stock Price Prediction Model Based on Investor Sentiment and Optimized Deep Learning	2023	#19
[26]	Accurate Stock Price Forecasting Based on Deep Learning and Hierarchical Frequency Decomposition	2024	#20
[27]	Artificial Intelligence in Accounting and Finance : Challenges and Opportunities	2023	#21
[28]	Decision Fusion for Stock Market Prediction : A Systematic Review	2022	#22
[29]	Forecasting Stock Prices Using a Hybrid Deep Learning Model Integrating Attention Mechanism, Multi- Layer Perceptron and Bidirectional Long-Short Term Memory Neural Network	2020	#23
[30]	How to Handle Data Imbalance and Feature Selection Problems in CNN-Based Stock Price Forecasting	2022	#24
[31]	Hybrid Information Mixing Module for Stock Movement Prediction	2023	#25
[32]	Improving Financial Time Series Prediction Accuracy Using Ensemble Empirical Mode Decomposition and Recurrent Neural Networks	2020	#26
[33]	Improving Stock Price Prediction Using Combining Forecasts Methods	2012	#27
[34]	Multi-Granularity Spatio-Temporal Correlation Networks for Stock Trend Prediction	2024	#28
[10]	ML-GAT :A Multilevel Graph Attention Model for Stock Prediction	2022	#29
[35]	Transformer-Gated Recurrent Unit Method for Predicting Stock Price Based on News Sentiments and Technical Indicators	2023	#30
[1]	Multi-Layer Coupled Hidden Markov Model for Cross-Market Behavior Analysis and Trend Forecasting	2019	#31
[36]	Stock Prediction Based on Genetic Algorithm Feature Selection and Long Short-Term Memory Neural Network	2020	#32

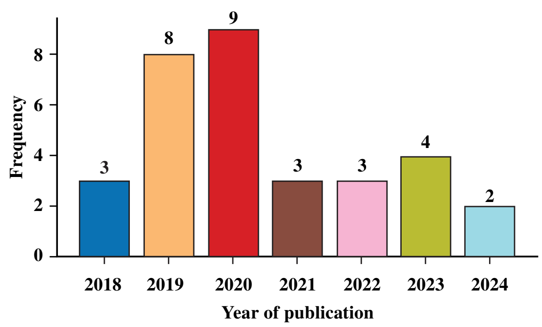

Fig. 1: Publication year of the studied papers.

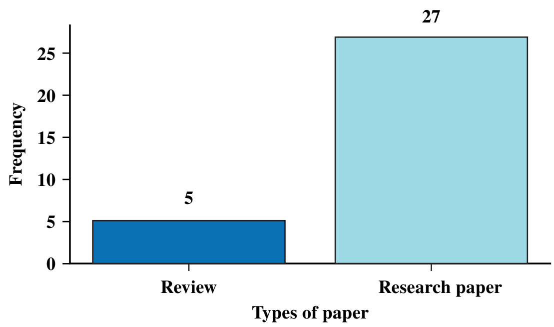

Fig. 2: Types of papers included in the study.

Gandhmal and Kumar[24] surveyed techniques used in stock prediction, categorizing them into classification and prediction methods. Analyzing 50 papers, they found that artificial neural networks (ANN) and fuzzy-based techniques are the most commonly employed. They also concluded that incorporating multiple factors is essential for accurate and efficient market prediction.

Table 3: Synthetic table of research articles-part 1

Ref	Technique	Based models	Type of Metrics	Type of Inputs	Historical Dates	Data Sources	Underlying	Output
[15]	Hybrid	IGCN; Dual-CNN	Accuracy, Preci-sion, Recall, F1-score, AAR, SR, TMoney	Stock correlations, Market data, Technical indicators	2013/01/01–2020/01/01; 2015/07/01–2020/01/01; 2012/01/01–2020/01/01; 2010/01/01–2020/01/01; 2013/01/01–2020/01/01	Shanghai Stock Exchange	Stock codes: 600809, 300330, 603808, 002580, 601318, 603123	J_J trend
[9]	Hybrid	NLP; RNN-LSTM	Accuracy, Loss	Sentiment index	2003–2013	Thomson Reuters News Archive; Thomson Reuters Tick History	Dow Jones 30 index	J_J trend
[16]	Hybrid	MDTW; CNN	Accuracy, Recall, Precision, F1-score	Market data	02/01/1962–15/09/2017	Yahoo Finance	S&P 500 index; GOOGL, BA, IBM, DIS, GT, APPL, GE	J_J trend
[17]	Hybrid	Knowledge graph; Graph embedding; CNN; BiLSTM	Accuracy, Balanced accuracy, AUC	Stock market information, Trading data	03/2012–06/2018	A top-five Chinese securi-ties company	CITIC Securities, China Pin-gan	J_{J+1} trend
[3]	Machine Learning	Transfer entropy; MKL	Accuracy, F1-score, p-value	Causal relationships, Mar-ket data	01/01/2014–31/12/2016	Naver; KOSCOM	GICS sectors: Pharmaceuti-cals, Materials, Food Expenses	J_J trend
[8]	Decision Fusion	NNRE, SVRE, BRT, RFR	RMSE, MAPE, MAE	Market data, Technical in-dicators, Sentiment index, Search volume, Trend data	01/01/2013–31/12/2016	Yahoo Finance; R Package; Quandl; Google other stocks Trends; Wikimedia	Citigroup stock; 19 (Facebook, McDonald's, . . .)	J_{J+1} price
[19]	Hybrid	EMD; FNN	RMSE, MAPE, MAE, SR, TR, AAR, MD, AAR/MD	Market data	04/01/2012–30/12/2016; 03/01/2007–30/12/2011	Yahoo Finance	SSEC index, Nasdaq index, S&P 500 index	J_J price
[20]	Hybrid	CDM; LSTM; CEEMD; CNN; LSTM	RMSE, MAE, MAPE	Market data	01/2010–10/2019	Yahoo Finance	S&P 500 index, Dow Jones index, DAX index, Nikkei 225 index	J_{J+1} price
[21]	Deep Learn-ing	ANN, RF	RMSE, MAPE,	Market data MBE	04/05/2009–04/05/2019	Yahoo Finance	Nike; Goldman Sachs; John-son & Johnson; Pfizer; JP Morgan Chase & Co	J_{J+1} price
[7]	Decision Fusion	RF, ERT, RNN, B-RNN, RNN-LSTM, GRU, XGBoost, LightGBM	Accuracy, F1-score, AUC	Technical and macroeco-nomic indicators	01/2003–04/2009	Yahoo Finance; Federal Reserve Bank of St. Louis	S&P 500 index, Dow Jones index, Nasdaq index trend	$JJ+20$

Table 4: Synthetic table of research articles-part 2

Ref	Technique	Based Models	Type of Metrics	Type of Inputs	Historical Dates	Data Sources	Underlying	Output
[2]	Hybrid	ARFIMA; LSTM	RMSE, MSE, MAPE	Market data	01/01/2009–30/05/2018	Islamabad Pakistan Stock Exchange	Fauji Fertilizer Company	J_J price
[22]	Hybrid	LDA-POS; SVM	Accuracy	Social media data; Market data	31/07/2012–19/07/2013; 30/04/2016–13/09/2016	Yahoo Message Board; Yahoo Finance; SAHAMYAB; Tehran Securities Exchange	18 U.S. stocks; 5 Iranian stocks	J_J trend
[4]	Machine Learning	Coupled matrix; Tensor factorization	Accuracy, MCC	Market data; Social media data; News articles	01/01/2015–31/12/2015	China A-share; HK Stock Market; Guba; Xueqiu; Wind	78 Chinese stocks; 13 Hong Kong stocks	J_J trend
[5]	Deep Learning	LSTM	RMSE	Technical indicators; Mar- ket data	01/09/2017–16/02/2018	Kaggle	S&P 500 index; Amazon	1-min price; 5- min price; 10-min price
[6]	Deep Learning	Associated Network	MSE, MAE, Average Accuracy	Technical indicators	Not mentioned	Shanghai Stock Exchange; Shenzhen Stock Exchange	SSEC index; PetroChina; ZTE	J_{J+1} price
[25]	Hybrid	LSTM; SSA; Sentiment Analysis	MAPE, MAE, RMSE, R^2	Technical indicators; Sen- timent index	01/07/2016–30/06/2022	East Money Net; Ruisi Financial Database	PetroChina; CITIC Securities; Guizhou Bailing; HiFuture Technology; Xinning Logistics; Zhongke Electric	J_J price
[26]	Hybrid	CEEMDAN; K-means; VMD; GRU	MAPE, MAE, RMSE, R^2	Market data	19/12/1990–25/05/2023; 03/04/1991–25/05/2023; 03/12/1990–25/05/2023	Yahoo Finance	SSEC index; SZI; S&P 500 index	J_J price
[29]	Hybrid	AM; MLP; BiL- STM	MSLE, MedAE, EVS, MSE, MAE, R^2	Market data; Technical indicators; Natural resource market data; Google search indices	02/09/2008–12/07/2019	Yahoo Finance	S&P 500 index; Nasdaq index; Russell 2000 index; Dow Jones index	J_J price
[30]	Deep Learning	CNN	Accuracy, Recall, Precision, F1-score tocks	Technical indicators; Natural resource market data (gold and oil prices)	01/01/2008–15/12/2021	Yahoo Finance	Dow Jones index; 10 other	J_J price
[31]	Hybrid	GRU; BERT; MLP; GAP	Accuracy, Precision, Recall, Specificity, F1- score, MCC	Market data; Tweets	01/01/2014–01/01/2016	StockNet database	85 stocks from the S&P 500 index	J_{J+1} trend

Table 5: Synthetic Table of Research Articles-Part 3

Ref	Technique	Based Models	Type of Metrics	Type of Inputs	Historical Dates	Data Sources	Underlying	Output
[32]	Hybrid	EMD; Sample Entropy; LSTM	MAPE, WAPE, DA, TheilU,	Market data	01/02/2018–17/01/2020	Yahoo Finance	AT&T Company; The Boeing Company; 500 other stocks	J_J+1 price
[33]	Hybrid	EMD; ARIMA; EWMA	MAE, MAPE, RMSE	Market data	31/12/2013–21/08/2019	Not mentioned	Bank of Montreal; Bank of Nova Scotia; Imperial Oil Limited; Sun Life Financial Inc.; TC Energy Corporation	J_J price
[34]	Deep Learning	MLP; GRU	Accuracy, Precision, Recall, F1-score, CR, SR, ARR, MDD	Technical indicators; Features computed based on market data	28/09/2010–09/05/2023	Not mentioned	CSI 300 index; Nasdaq index	J_J trend
[10]	Hybrid ARV	LSTM; BERT; GNN	Accuracy, F1-score, Sharpe	Market data; Scientific articles; Stock correlation	08/02/2013–29/08/2019	Wikipedia company relationships; Yahoo Finance; Shanghai Stock Exchange; Accounting Research Database	423 stocks from the S&P 500 index; 286 stocks from the CSI 300 index	J_J trend
[35]	Hybrid	BERT; TEGRU	Accuracy, F1-score, AcMAPE, MSE, MAPE, R2, RMSE	Sentiment aspect scores; Technical indicators	01/01/2017–31/03/2022	Indonesian news portal	5 highest-capitalization (kontan.co.id); Yahoo Finance	J_J price Indonesian with the largest gains
[1]	Statistical	MCHMM (Multi-Layer Coupled Hidden Markov Model)	Accuracy, Precision, Type I error, ROR, AAR	Market data	01/2008–12/2018	Economic Research	Dow Jones index	J_J trend
[36]	Hybrid	Genetic	MSE	Market data; Technical indicators; Financial indicators	01/01/2010–01/04/2020	JoinQuant platform	Quantitative	China Construction 300 index

Table 6: Summary of reviewed papers

Reference	No. of reviewed papers	Priced asset class	Employed technique	Most used models
[18]	86	Forex; Stock	Deep learning	CNN; LSTM; DNN; RNN; Reinforcement learning; HAN; NLP; Wavenet
[24]	50	Stock	Machine learning; Deep learning	ANN, SVM, SVR, HMM, NN, fuzzy
[27]	21	Stock; Option	Deep learning; Machine learning	For stock pricing: SVM, ANN, Text mining, Data mining; For option pricing: intelligent algorithms, Machine learning, Deep learning
[28]	75	Stock	Decision fusion	Homogeneous base learners: ANN, decision trees, SVM, LSTM; Heterogeneous base learners employing multiple algorithms
[5]	32	Stock	Statistical, Pattern recognition, Machine learning, Deep learning, Sentiment analysis, Hybrid	For statistical methods: ARIMA, ESN, Regression ...; Pattern recognition: template matching, PIP ...; Machine learning: ANN, SVR, XGBoost ...; Sentiment analysis: Text mining, Google Profile of Moods, Opinion Finder ...; Deep learning: LSTM, RNN ...

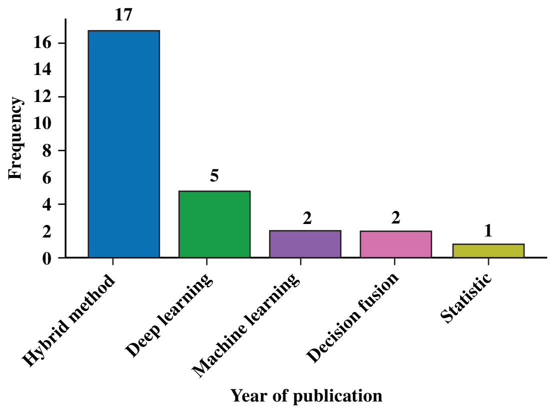

Fig. 3: Prediction techniques used in research articles

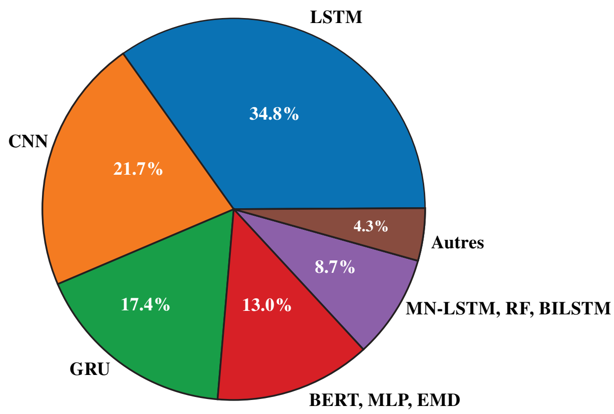

Fig. 4: Model types used across prediction techniques

Yi et al.[27] reviewed the application of artificial intelligence to classical problems in accounting and finance, demonstrating its use in financial statement analysis, corporate valuation, financial asset pricing, risk management, and so on. In the context of financial asset pricing, AI is mainly applied to stock and option pricing.

Zhang et al.[28] conducted a systematic review of decision fusion methods in stock prediction. They reported that decision fusion models consist of either homogeneous base learners, such as ANN, decision trees, SVM and LSTM, or heterogeneous ensembles that integrate multiple algorithms. They also noted that the use of identical heterogeneous learner sets across studies is uncommon.

Shah et al.[5] reviewed stock market prediction techniques and proposed a taxonomy of methods, while also discussing challenges and opportunities in stock analysis and forecasting.

Table 6 summarizes these review papers. Across these studies, several asset classes are examined, including stocks, options and foreign exchange. The techniques employed encompass statistical methods, pattern recognition, machine learning, deep learning, decision fusion, sentiment analysis and hybrid approaches.

Analysis of research articles

Prediction techniques used in research articles: Figure 3 illustrates the techniques used for predicting stock and index prices in the research articles. These techniques can be classified into five categories:

- Hybrid techniques: 63% of research articles
- Deep learning techniques: 19%
- Machine learning techniques: 7%
- Decision fusion techniques: 7%
- Statistical techniques: 4%

Hybrid techniques, which integrate multiple methods, are the most prevalent, accounting for 63% of the research articles. This finding indicates a growing interest in hybrid approaches in recent years. These categories largely correspond to those identified in the review articles, with the exception of pattern recognition techniques.

Model types used by prediction techniques: Figure 4 presents the most frequently used models across the different prediction techniques. A total of 51 models are employed across all research articles. Deep learning models (e.g., LSTM, CNN, GRU) are the most prevalent, followed

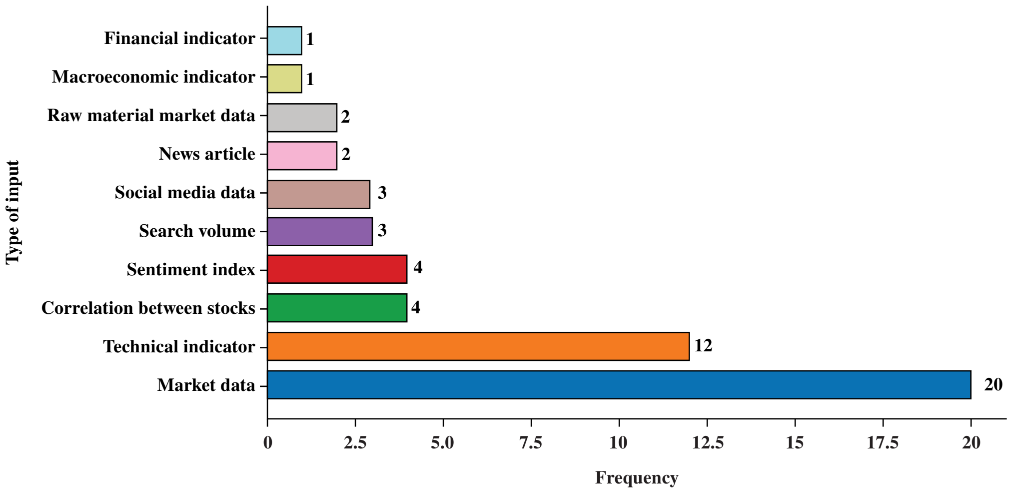

Fig. 5: Types of inputs used across techniques and models

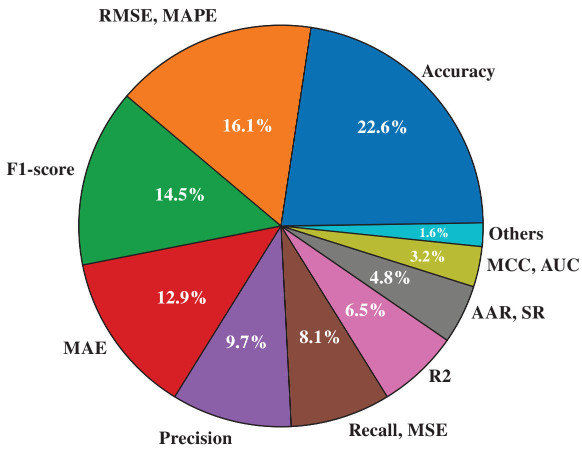

Fig. 6: Evaluation metrics used across techniques

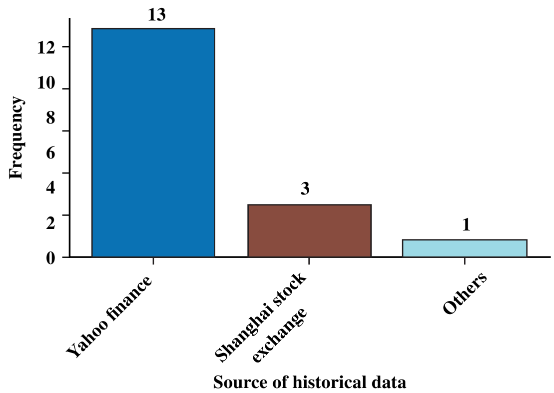

Fig. 7: Sources of extracted historical data

by classical machine learning models (e.g., MLP, RF), statistical models (e.g., ARIMA) and signal decomposition models (e.g., EMD). The nine most frequently used models are highlighted in the figure.

Input types used for model training: Figure 5 illustrates the types of inputs used to train the models. Ten input categories were identified, ranked by frequency: Trading data (historical prices), technical indicators, stock correlations, sentiment indices, search volume, social media data, news articles, raw material market data, macroeconomic indicators and financial indicators.

Trading data is the most widely used input (20 articles), followed by technical indicators (12 articles). All other input types appear in fewer than five articles each.

Evaluation metrics used for machine learning models: Model performance is assessed using two main approaches: evaluation on a holdout test dataset and performance assessment within a trading scenario using investment strategies. Figure 6 presents the evaluation metrics employed. A total of 41 metrics are reported, with the 13 most common shown in the figure. Accuracy, RMSE, MAPE, MAE and Precision are among the most frequently used metrics.

Historical data sources and temporal windows: Figure 7 shows the data sources used to obtain historical data. Across the articles, 33 distinct sources are reported. Yahoo Finance is the most frequently used source (13 occurrences, 40%), followed by the Shanghai Stock Exchange (3 occurrences, 9%). All other sources are used once each (3%).

Most studies rely on historical data ranges of fewer than 5,000 days. Some articles use a single data range, whereas others employ multiple ranges. The largest datasets were analyzed by Wen et al.[16] (>20,000 days), followed by Li et al.[26] (>12,000 days) and Akşehir and Kılıç[30] (5,000 days).

Underlying financial instruments used for prediction: Figure 8 presents the financial instruments used as underlyings for prediction. A total of 57 underlyings are considered. Indices are used more frequently than individual stocks, with the S&P 500 being the most common (8 articles), followed by the Dow Jones (5) and Nasdaq (4).

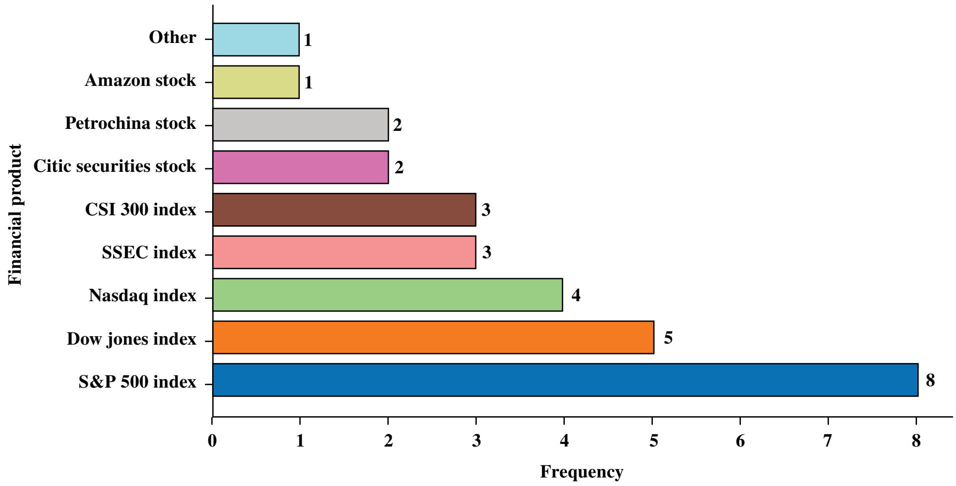

Fig. 8: Underlyings of stocks and indices used for prediction.

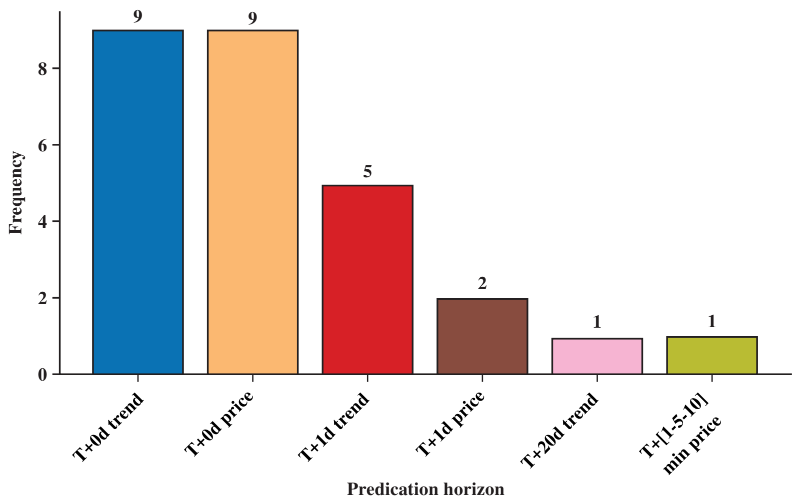

Fig. 9: Type and horizon of financial prediction.

Other indices include SSEC and CSI 300 (3 occurrences each). Among individual stocks, CITIC Securities and PetroChina are the most frequently used (2 occurrences each). Most underlyings appear only once.

Type and prediction horizon: Two main types of predictions are identified: Price prediction and trend prediction, over various time horizons. Figure 9 illustrates the distribution of prediction horizons. Most studies(18) focus on same-day predictions, including 9 for price and 9 for trend. Five articles predict next-day prices and two predict next-day trends. One article predicts trends 20 days ahead, while another predicts three prices simultaneously (after 1, 5 and 10 minutes). Overall, prediction horizons are predominantly short term. Future research could extend this work by exploring longer-term predictions of stock and index prices or trends.

CONCLUSION

This study presented a systematic literature review of techniques used for predicting financial assets, specifically stocks and indices, from 2018 onward. The PRISMA methodology was followed, involving a structured multi-stage process. Relevant articles were retrieved from Scopus and IEEE Xplore and after applying the inclusion and exclusion criteria, 32 articles were retained, comprising 5 review papers and 27 research articles.

Several key findings emerged from the synthesis of these studies. The review papers highlight a wide range of asset classes and methodological approaches in financial prediction, including statistical methods, pattern recognition, machine learning, deep learning, decision fusion, sentiment analysis and hybrid techniques.

Analysis of the research articles indicates that prediction techniques can be categorized into five main groups: Hybrid, deep learning, machine learning, ensemble and statistical approaches. Hybrid techniques, which integrate multiple methods, are the most prevalent, appearing in 17 studies and accounting for approximately 63% of the research articles.

Across these techniques, a variety of model types are employed, including deep learning, classical machine learning, statistical and signal decomposition models. Deep learning models dominate, with LSTM, CNN and GRU being the most frequently used architectures.

Model performance is typically evaluated using two main strategies: testing on a holdout dataset and assessing performance within simulated trading scenarios. A wide range of evaluation metrics is applied, with Accuracy, RMSE, MAPE and F1-score being the most commonly reported.

Ten distinct types of input data are used across models and techniques, with trading data-primarily historical closing prices of stocks and indices-being the most frequently employed. These data are sourced from various platforms, with Yahoo Finance emerging as the most commonly used source.

Finally, the reviewed techniques address two primary prediction tasks: price prediction and trend prediction. The prediction horizons are predominantly short term, indicating that most studies focus on near-term forecasting.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the colleagues and peers who provided valuable feedback and constructive suggestions during the preparation of this manuscript.

REFERENCES

[1] W. Cao, W. Zhu and Y. Demazeau, "Multi-layer coupled hidden markov model for cross-market behavior analysis and trend forecasting," *IEEE Access*, vol. 7, pp. 158563-158574. 2019.

[2] A. H. Bukhari, M. A. Z. Raja, M. Sulaiman, S. Islam, M. Shoaib and P. Kumam, "Fractional neuro-sequential ARFIMA-LSTM for financial market forecasting," *IEEE Access*, vol. 8, pp. 71326-71338. 2020.

[3] K. Nam and N. Seong, "Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market," *Decis Support Syst.*, vol. 117, pp. 100-112. 2019.

[4] X. Zhang, Y. Zhang, S. Wang, Y. Yao, B. Fang and P. S. Yu, "Improving stock market prediction via heterogeneous information fusion," *Knowledge-based Syst.*, vol. 143, pp. 236-247. 2018.

[5] D. Shah, H. Isah and F. Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," *Int. J. Financ Stud.*, vol. 7, no. 2, 2019. [Online]. Available: 10.3390/ijfs7020026

[6] G. Ding and L. Qin, "Study on the prediction of stock price based on the associated network model of LSTM," *Int. J. Mach. Learn. Cybern*, vol. 11, pp. 1307-1317. 2019.

[7] M. Jiang, J. Liu, L. Zhang and C. Liu, "An improved stacking framework for stock index prediction by leveraging tree-based ensemble models and deep learning algorithms," *Physica A: Stat Mech Appl.*, vol. 541, 2020. [Online]. Available: 10.1016/j.physa.2019.122272

[8] B. Weng, L. Lu, X. Wang, F. M. Megahed and W. Martinez, "Predicting short-term stock prices using ensemble methods and online data sources," *Expert Syst. Appl.*, vol. 112, pp. 258-273. 2018.

[9] W. Souma, I. Vodenska and H. Aoyama, "Enhanced news sentiment analysis using deep learning methods," *J. Comput. Social Sci.*, vol. 2, pp. 33-46. 2019.

[10] K. Huang, X. Li, F. Liu, X. Yang and W. Yu, "ML-GAT:A multilevel graph attention model for stock prediction," *IEEE Access*, vol. 10, pp. 86408-86422. 2022.

[11] P. Soni, Y. Tewari and D. Krishnan, "Machine learning approaches in stock price prediction: A systematic review," in *Proceeding of 1st International Conference on Artificial Intelligence, Computational Electronics and Communication System (AICECS 2021)*, Oct. 2021. Manipal, India: IOP Publishing, 2022, pp. 1-10.

[12] L. N. Mintarya, J. N. M. Halim, C. Angie, S. Achmad and A. Kurniawan, "Machine learning approaches in stock market prediction: A systematic literature review," *Procedia Comput. Sci.*, vol. 216, pp. 96-102. 2023.

[13] D. Kumar, P. K. Sarangi and R. Verma, "A systematic review of stock market prediction using machine learning and statistical techniques," *Mater. Today: Proc.*, vol. 49, no. 8, pp. 3187-3191. 2022.

[14] C. Y. Lin and J. A. L. Marques, "Stock market prediction using artificial intelligence: A systematic review of systematic reviews," *Social Sci. Hum. Open*, vol. 9, 2024. [Online]. Available: 10.1016/j.ssaho.2024.100864

[15] W. Chen, M. Jiang, W.-G. Zhang and Z. Chen, "A novel graph convolutional feature based convolutional neural network for stock trend prediction," *Information Sci. s*, vol. 556, 2021. [Online]. Available: 10.1016/j.ins.2020.12.068

[16] M. Wen, P. Li, L. Zhang and Y. Chen, "Stock market trend prediction using high-order information of time series," *IEEE Access*, vol. 7, pp. 28299-28308. 2019.

[17] J. Long, Z. Chen, W. He, T. Wu and J. Ren, "An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in chinese stock exchange market," *Appl. Soft Comput.*, vol. 91, 2020. [Online]. Available: 10.1016/j.asoc.2020.106205

[18] Z. Hu, Y. Zhao and M. Khushi, "A survey of forex and stock price prediction using deep learning," *Appl. Syst. Innovation*, vol. 4, no. 1, 2021. [Online]. Available: 10.3390/asi4010009

[19] F. Zhou, H.-m. Zhou, Z. Yang and L. Yang, "EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction," *Expert Syst. Appl.*, vol. 115, pp. 136-151. 2019.

[20] H. Rezaei, H. Faaljou and G. Mansourfar, "Stock price prediction using deep learning and frequency decomposition," *Expert Syst. Appl.*, vol. 169, 2021. [Online]. Available: 10.1016/j.eswa.2020.114332

[21] M. Vijh, D. Chandola, V. A. Tikkiwal and A. Kumar, "Stock closing price prediction using machine learning techniques," *Procedia Comput. Sci.*, vol. 167, pp. 599-606. 2020.

[22] A. Derakhshan and H. Beigy, "Sentiment analysis on stock social media for stock price movement prediction," *Eng. Appl. Artificial Intell.*, vol. 85, pp. 569-578. 2019.

[23] Z. Lanbouri and S. Achchab, "Stock market prediction on high frequency data using long-short term memory," *Procedia Comput. Sci.*, vol. 175, pp. 603-608. 2020.

[24] D. P. Gandhmal and K. Kumar, "Systematic analysis and review of stock market prediction techniques," *Comput. Sci. Rev.*, vol. 34, 2019. [Online]. Available: 10.1016/j.cosrev.2019.08.001

[25] G. Mu, N. Gao, Y. Wang and L. Dai, "A stock price prediction model based on investor sentiment and optimized deep learning," *IEEE Access*, vol. 11, pp. 51353-51367. 2023.

[26] Y. Li, L. Chen, C. Sun, G. Liu, C. Chen and Y. Zhang, "Accurate stock price forecasting based on deep learning and hierarchical frequency decomposition," *IEEE Access*, vol. 12, pp. 49878-49894. 2024.

[27] Z. Yi, X. Cao, Z. Chen and S. Li, "Artificial intelligence in accounting and finance: Challenges and opportunities," *IEEE Access*, vol. 11, pp. 129100-129123. 2023.

[28] Z. Yi, X. Cao, Z. Chen and S. Li, "Artificial intelligence in accounting and finance: Challenges and opportunities," *IEEE Access*, vol. 11, pp. 129100-129123. 2023.

[29] Q. Chen, W. Zhang and Y. Lou, "Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron and bidirectional long-short term memory neural network," *IEEE Access*, vol. 8, pp. 117365-117376. 2020.

[30] Z. D. Akşehir and E. Kiliç, "How to handle data imbalance and feature selection problems in cnn-based stock price forecasting," *IEEE Access*, vol. 10, pp. 31297-31305. 2022.

[31] J. Choi, S. Yoo, X. Zhou and Y. Kim, "Hybrid information mixing module for stock movement prediction," *IEEE Access*, vol. 11, pp. 28781-28790. 2023.

[32] H. D. Chacon, E. Kesici and P. Najafirad, "Improving financial time series prediction accuracy using ensemble empirical mode decomposition and recurrent neural networks," *IEEE Access*, vol. 8, pp. 117133-117145. 2020.

[33] M. R. Hossain, M. T. Ismail and S. A. B. A. Karim, "Improving stock price prediction using combining forecasts methods," *IEEE Access*, vol. 9, pp. 132319-132328. 2021.

[34] J. Chen, L. Xie, W. Lin, Y. Wu and H. Xu, "Multi-granularity spatio-temporal correlation networks for stock trend prediction," *IEEE Access*, vol. 12, pp. 67219-67232. 2024.

[35] A. T. Haryono, R. Sarno and K. R. Sungkono, "Transformer-gated recurrent unit method for predicting stock price based on news sentiments and technical indicators," *IEEE Access*, vol. 11, pp. 77132-77146. 2023.

[36] S. Chen and C. Zhou, "Stock prediction based on genetic algorithm feature selection and long short-term memory neural network," *IEEE Access*, vol. 9, pp. 9066-9072. 2021.