Journal of Theoretical and Applied Information Technology

15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEEP LEARNING IN STOCK MARKET PREDICTION: A FIVE-YEAR LITERATURE REVIEW ON DEVELOPMENTS, CHALLENGES, AND FUTURE DIRECTIONS

BERNADECTUS YUDI DWIANDIYANTA[1] , RUDY HARTANTO[2], RIDI FERDIANA[3]

[1,2,3]Universitas Gadjah Mada, Department of Electrical Engineering and Information, Indonesia

[1]Universitas Atma Jaya Yogyakarta, Department of Informatics, Indonesia

E-mail: [1]yudi.dwiandiyanta@uajy.ac.id, [2]rudy@ugm.ac.id, [3]ridi@ugm.ac.id

ABSTRACT

Shares or equities have received significant attention in investment because of their profit potential. However, with the complexity and volatility in the stock market, the need arises for more accurate prediction methods. In the last decade, Deep Learning algorithms have become a promising solution. Deep learning methods offer superior capabilities in handling big data and non-linear relationships. This research reviews the development of Deep Learning algorithms in stock market predictions from 2017 to 2022. This research uses a literature review methodology. This research reviewed 86 articles selected from the initial 346 articles for further analysis. The analysis results show the dominance of using historical trading data as system input in stock price predictions. The New York and Shanghai Stock Exchanges are the main focus of researchers' attention. This research also identifies the potential for combining input data and optimizing prediction methods as a future research opportunity. This research generates opportunities to develop algorithms based on LSTM and Bi-LSTM, hybrid and ensemble methods. It is hoped that this research can provide insight into the latest developments in stock prediction using Deep Learning and provide insight into future research directions.

Keywords: *Stock Market Prediction, Deep Learning, Literature Review, Feature Engineering, Stock Forecasting*

1. INTRODUCTION

Stock, often referred to as equity, represents a form of investment that signifies ownership within the issuing firm [1]. Stock is mostly bought and sold on stock market and serve as the foundation of the portfolios of many retail investors. Stock is a popular investment option among many investors due to their ability to generate a high amount of profit. When investors hold shares, they have two benefits, including benefiting from growing stock prices (Capital Gain) and receiving business dividends [2][3]. Subsequently, the capital gain is the profit received from the difference between the selling price of stock and the purchase price of the shares. In contrast, the dividend is the profit distribution supplied by the firm and is derived from the firm's earnings [2][3].

To maximize the profitability of stock investments, investors typically employ two types of analysis, which are fundamental and technical analysis [4]. The primary goal of these analytical methods is to enhance the returns on their investments. Subsequently, the fundamental analysis evaluates stock based on economic, industry, and overall company conditions. While technical analysis employs historical stock trading data, transaction volumes, stock chart patterns, seasonal analysis, and other methods. Technical analysis operates under the premise that it is feasible to anticipate future price movements within market by scrutinizing historical price data.

Over the course of time, the exploration of financial time series analysis and prediction of stock prices has evolved into a captivating research domain. Stock price movements that are influenced by many factors make investors able to receive uncertain returns and potential losses. Consequently, prediction of stock prices and the selection of investment-worthy stock have emerged as pivotal concerns for investors. Due to the complexity and volatility of stock prices [5], a large number of factors and sources of information must be taken into account when trying to forecast stock prices. This is a very challenging endeavor and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

continues to be a topic of discussion in the financial market industry [6].

Deep Learning algorithms have showed superior performance compared to other methodologies in the stock price prediction. These algorithms exhibit the capability to handle substantial volumes of data and decipher non-linear associations between input variables and predictive targets. Such an advantage is not found in predictive models using linear or machine learning algorithms. Numerous research have examined Deep Learning applications in financial market [7][8][9]. Among these investigations, the topic of research revolves around Decision Fusion for Stock Market forecasting [10]. Shah et al. present numerous strategies for forecasting stock market that integrate Auto-Regressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), Hybrid LSTM, Convolutional Neural Network (CNN), and Hybrid CNN, as well as the drawback and precision of the many models [11].

This study commences with distinct objectives. Primarily, it endeavors to offer an exhaustive review of the advancements in the application of Deep Learning algorithms for stock market prediction spanning from 2017 to 2022. Furthermore, the study aims to discern the prevailing methods, pinpoint the types of input data predominantly employed, and identify the stock exchanges that have consistently garnered substantial scholarly attention. A central objective also lies in acknowledging the challenges that researchers confront in this dynamic domain. Conclusively, guided by discernible lacunae in existing literature, this research aspires to suggest prospective avenues for future scholarly exploration.

Research in this field has experienced many extraordinary developments, making it difficult for novice researchers to keep up with the latest research developments. In this research, the development of Deep Learning algorithms for stock market prediction is presented. Furthermore, the development of the method over the past five years is the focus, and the trends for each predictive model workflow that allows budding researchers to learn outdated technologies are provided.

This research focuses on comparing Deep Learning research in stock forecasting and predicting area. Additionally, particular attention is given to the execution and repeatability of prior research, aspects often overlooked in comparable surveys. In this literature review, based on a summary of the papers evaluated, the aim is to propose future research directions that can assist the reader in determining their next course of action. Comparisons are made regarding input features and output characteristics, dataset characteristics, and methods used. The aspiration is that through this literature review, the development of Deep Learning algorithms for attaining more precise prediction outcomes becomes feasible.

The primary contributions to this literature review are as follows:

(1) This research examines the latest developments in the use of Deep Learning algorithms for stock market prediction focusing on research that has emerged during the past five years.
(2) Based on a generic procedure for stock market forecasting, past investigations may be simply categorized and summarized. In each step of the workflow, subsequent research might refer to prior work.
(3) Specific analyses are offered in certain research, aspects that are frequently disregarded in comparable research.

Numerous studies have delved into the realm of stock prediction using deep learning methodologies. Historically, these investigations predominantly emphasized the technical aspects, focusing on the efficiencies of different algorithms and input parameters. The literature is replete with models attempting to optimize stock predictions through intricate manipulations of neural networks, LSTM structures, and other deep learning architectures. However, a conspicuous lacuna is evident in these prior works [12][13]: a comprehensive survey detailing the myriad factors influencing stock prediction models, from input parameters to dataset characteristics, remains largely uncharted. Our study seeks to bridge this gap. Contrary to the prevailing literature, which predominantly focuses on specific prediction models, our study offers a comprehensive overview, systematically categorizing and evaluating the diverse components utilized in the field over the past half-decade. This integrative approach not only elucidates the contemporary state of the art but also highlights prospective pathways for subsequent research.

This article is organized as follows: The first section presents an introduction. The second section presents the methodology used in this research. The third section contains the results. Finally, the fourth section presents the conclusions and future works.

Journal of Theoretical and Applied Information Technology
15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. METHODOLOGY

2.1 Criteria for Problem Selection

Prior to the initiation of the literature review, it is imperative to delineate the criteria underpinning the decision to concentrate on the prediction of stock prices via Deep Learning methodologies:

(1) Economic Relevance: The ability to accurately predict stock prices is of paramount importance, given its profound implications for formulating investment strategies and tactical maneuvers within the stock market ecosystem.

(2) Technological Advancements in Deep Learning: Recent strides in the domain of Deep Learning offer promising avenues for enhancing both the precision and computational efficiency of stock price prognostications.

(3) Proliferation of Data: The ubiquitous availability of historical stock datasets affords a conducive environment for the efficacious application of Deep Learning paradigms.

2.2 Review Methods

Over the past few years, numerous deep-learning model articles on stock prediction have been published. This research adopts an method based on the methodologies outlined in [14][15] to conduct the most feasible literature review. Even though the focus of these articles is software engineering, their methodologies are general and have been used in stock market literature [16].

Literature Review consists of three key stages: planning, execution, and reporting. The initial step (Step 1) involves defining the prerequisites for literature review. The objectives of conducting literature research was outlined at the beginning of this chapter. Subsequently, the current literature review of stock prediction are found and assessed. The methodology was designed to remove the possibility of researcher bias and direct the implementation of review (Step 2). This encompassed formulating research topics, search methods, research selection with inclusion and exclusion criteria, quality evaluation, data extraction, and synthesis process. Review process was created, reviewed, and iteratively modified during review's conducting and reporting phases.

2.3 Search Strategy and Research Selection

In this research, the articles analyzed were collected using tools provided at www.watase.web.id. Initially, the keywords used are "Stock AND Prediction AND Deep AND Learning" OR "stock AND forecasting AND deep AND learning." These keywords were deemed essential to ensure the alignment of reviewed papers with the designated subject matter. After this initial phase, a total of 346 journal articles were initially retrieved. Subsequently, a filtering process was enacted by employing the following criteria: only journal articles indexed in Scopus and published within the timeframe of 2017 to 2022 were considered suitable for inclusion in this article. Following this filtering, a refined list of 346 journal articles emerged. The subsequent stage involved a manual screening process. Here, the focus was directed towards identifying solely those journal articles that closely pertained to stock prediction through the application of Deep Learning methods. These selected articles will serve as the foundation for referencing within this research. The selection process was meticulously guided by predetermined inclusion and exclusion criteria, resulting in the identification of the primary research that were aligned with the established criteria. The following inclusion and exclusion criteria are shown in Table 1. As a result of this third phase, a collection of 86 journal articles was pinpointed. The article selection process is shown in Figure 1.

Table 1: Inclusion and Exclusion Criteria.

Exclusion Criteria	Papers that do not correspond to the intended subject
	Review paper and comparison paper
	Research without a strong validation
	Studies not written in English
	Studies with the topic of clustering
	Studies with object research focus on commodity, currency, and futures
Inclusion Criteria	Studies in Stock Exchange all around the world
	Studies using deep learning algorithm
	Studies in predicting market trends and predicting stock price

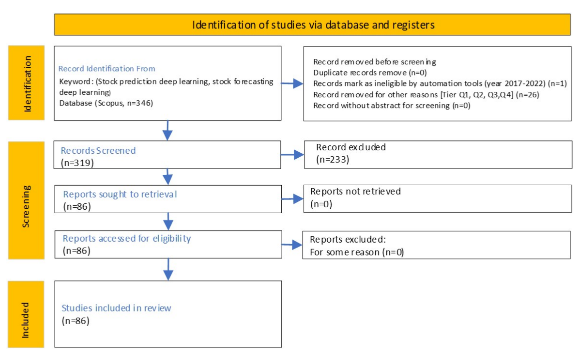

Figure 1: Article Selection Process

Journal of Theoretical and Applied Information Technology
15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. RESULT

Prior research examined the use of Deep Learning algorithms to forecast the price/movement of stock market using a variety of data source combinations. In this part, the primary research is characterized as a basic process consisting of five phases to which the majority of research adhere: input feature, output characteristics, dataset characteristics, and prediction methods. The subsequent part will throw more light on each of these steps separately and provide general methods that can be easily replicated in future projects.

3.1 The Input Feature

Feature engineering refers to the process of collecting useful information from unstructured data through the use of domain expertise to feed it into machine learning models. These hand-crafted features are inputs for prediction models and, when combined with raw data, may significantly enhance the effectiveness of machine learning models.

3.1.1 Stock Historical Data

Stock historical data entails information that encapsulates the evolution of individual stock prices, pertaining to issuers whose shares are actively traded on a country's stock exchange. This data is organized within specific time intervals. Stock trading data can be in the form of open, high, low, and closed prices at certain intervals. In several investigations, several researchers use Open, High, Low, Close, and Volume data, but some only use closing prices.

3.1.2 Technical Indicator

Technical Analysis employs indicators, which are mathematical formulas designed to assess prevailing market conditions and aid in generating buy or sell signals. Currently, there are hundreds or thousands of indicators that have been developed. Each indicator has its character and method of application. Table 2 shows the indicator that is used in the articles reviewed. Table 3 shows the technical indicators used in each research.

Table 2: List of Indicator Technical.

Symbol	Indicator	Symbol	Indicator
MA	Moving Average	MOM	Momentum Indicator
SMA	Simple Moving Average	DX	Directional Movement Index
EMA	Exponential Moving Average	A/D	Accumulation/Distribution
WMA	Weighted Moving Average	AROSC	Aroon Oscillator
HMA	Hull Moving Average	ACLB	Acceleration Bands
TEMA	Triple Exponential Moving Average	CMF	Chaikin Money Flow
KAMA	Kaufman's Adaptive Moving Average	PSAR	Parabolic SAR
MACD	Moving Average Convergence/Divergence	VMAP	Volume-Weighted Average Price
STCK%	Stochastic Oscillator %K	RSV	Reserve
STCD%	Stochastic Oscillator %D	TR	True Range
WILLR%	Williams %R	MI	Mass Index
IMI	Intraday Momentum Index	OBV	On Balance Volume
RSI	Relative Strength Index	OSC	Oscillator
MFI	Money Flow Index	CMO	Chande Momentum Oscillator
ROC	Price Rate Of Change	ULTOSC	Ultimate Oscillator
CCI	Commodity Channel Index	KDJ	KDJ line
FI	Force Index	BB	Bollinger Band
ATR	Average True Range	BIAS	Deviation Rate
EMV	Ease of Movement		

Table 3: The Technical Indicators Used in Each Study.

Research	Indicator Technical
[17]	SMA, STCK%, STCD%, IMI, RSI, MFI, ROC, CCI, FI, ATR, DX, EMA
[18]	SMA, WMA, STCK%, STCD%, MACD, RSI, WILLR%, CCI, MOM, A/D
[19]	MACD, KAMA, AROSC, Acceleration ACLB, stochastic oscillator, CMF, PSAR, ROC, MOM, VWAP
[20]	MA, STCK%, RSV
[21]	TR, ATR, MI, RSI
[22]	Dataset1: SMA, WMA, RSI, STCK%, STCD% Dataset2: SMA, EMA, MACD, OSC, OBV
[23]	MA, EMA, MACD, RSI, WILLR, MOM, CMO, ULTOSC, OBV, ADOSC
[24]	EMA, MACD
[25]	MA, WMA, MACD, RSI, WR, CCI
[26]	CCI, MACD, WR, CMO, EMA, HMA, SMA, WMA, TEMA, ROC
[27]	MA, MACD, RSI
[28]	MACD, RSI, KDJ, MFI
[29]	RSI, STCK%, STCD%, Stochastic Slow %D, WILLR%, MACD, ROC, CCI
[30]	Stochastic Oscillator, WILLR%, RSI
[31]	MA, RSI, MACD
[32]	OBV, MA, BIAS

3.1.3 Text Data

Text data pertains to textual content originating from individuals, encompassing sources like social media, news articles, and online search queries. These data, which are a sort of alternative data, are difficult to obtain and interpret, but they may give important information that market data does not. These text data may be subjected to sentiment analysis to provide a sentiment analysis that can be used for forecasting.

3.1.4 Macroeconomics Data

Macroeconomic data, such as core inflation, economic growth rate, and Gross Domestic Product (GDP), represent the economic condition of a country, region, or industry. These indicators affect stock market because they represent the general condition of a country's economy. The impact of macroeconomic data is far-reaching, contributing significantly to the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

trajectory of stock market movements, encompassing both upward trends and downturns.

3.1.5 Fundamental Data

Accounting data, such as asset values, Price to Book Value, is the most common type of quarterly fundamental data. This kind of fundamental data is disclosed every three months. Due to its infrequent updates and the publication date being different for every company, this fundamental data is rarely used in prediction with Deep Learning models.

The selection of input features holds significant sway over both the accuracy and computational demands of stock prediction outcomes. Some of the input features often used in stock prediction research are shown in Figure 2. Notably, a predominant method remains rooted in the use of historical data alone. Approximately 53% of researchers exclusively employ historical data as input for their predictive systems. Most of the research developed using historical data uses complete Open, High, Low, Close, and Volume data [33][34][35][36][37][38][22], but some only use Close price data [39][40][41][42][43][44]. Historical data is considered insufficient, hence, it needs to be combined with technical indicators, fundamental data, and text data, usually present in sentiment analysis.

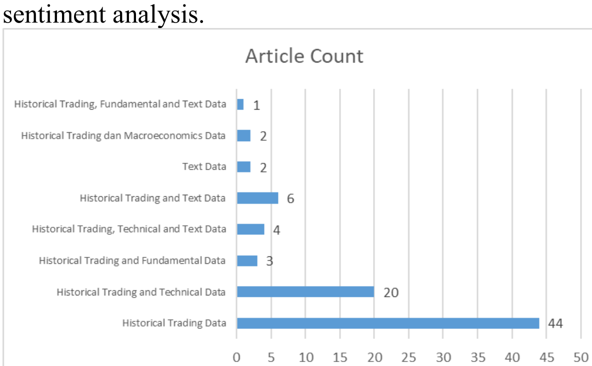

Figure 2: Input Feature Combination

Based on the papers reviewed, about 24.39% of the research uses historical data and technical indicators simultaneously. Liu et al. use a combination of historical data with Exponential Moving averages and MACD indicators [24]. Shilpa and Shambhavi use the indicators of the Moving Average, Relative Strength Index, and MACD [31]. Meanwhile, Kumar and Haider conducted research using many indicators, namely: SMA, Stochastic Oscillator, IMI, RSI, MFI, ROC, CCI, FI, ATR, EMV, DX, and EMA [17]. Thakkar et al. use SMA, WMA, STCK%, STCD%, MACD, RSI, WILLR%, CCI, MOM, and A/D Oscillator indicators to predict stock on India's National Stock Exchange (NSE) [18]. The combination of input features that are rarely examined is historical data and fundamental data [45][46]. Research using historical data, technical data, and text data in the form of news and sentiment was conducted by Chen [28], Wu et al. [30], and Shilpa and Shambhavi [31] received the attention of 4.88% of researchers. Meanwhile, research that gets the least attention from researchers is the investigation that involves historical data, fundamental data, and text data [47]. Figure 2 shows the number of articles based on the input feature combination method used.

3.2 Output Characteristics

The outcomes of stock prediction research can be categorized into three distinct groups: research with stock price prediction output, research with market direction (trends) prediction output, and both.

3.2.1 Type Prediction

3.2.1.1 Market Trends

Prediction of market trends has been carried out by many researchers. This line of inquiry revolves around predicting the direction in which market is moving, whether it's experiencing a bullish, bearish, or sideways trend. In addition, research is conducted to predict whether market is in crisis. Several research have also been developed to carry out investment strategies in stock market, namely the decision to buy, hold and sell. The algorithm used to predict market direction can be developed with an algorithm related to classification. Research related to market trends was conducted by [19][48][49][50][51]. Chatzis et al. examines predicting market crises using Deep Learning and machine learning. This research aims to predict whether a crisis occurs at a certain time [52].

3.2.1.2 Price Prediction

The predicted stock price is calculated using a certain method and then compared with the actual price that occurred. Research related to price prediction has been carried out by [53][54][55][56][57][58][59].

3.2.1.3 Both Market Trends and Price Prediction

Certain researchers undertake a dual method, delving into both prediction of market trends and the forecast of stock or index prices. Wang, Cheng, and Dong examine a multivariate deep-learning framework to predict stock index futures [60]. Li

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and Pan examine an ensemble deep-learning algorithm for stock prediction based on stock value and sentiment [61]. Figure 3 shows the number of studies in the three domains.

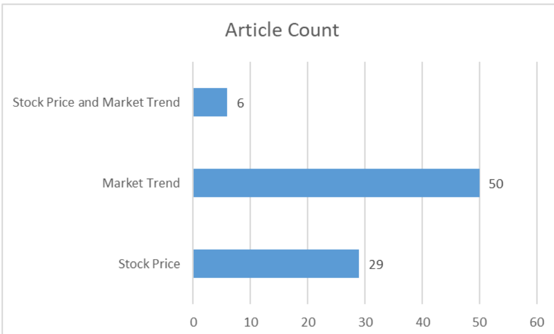

Figure 3: The Number of Studies in The Three Domains

3.2.2 Research Object

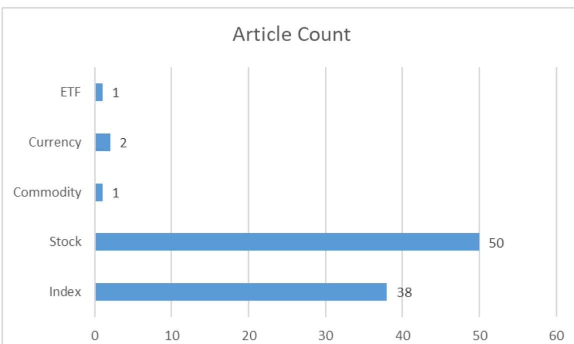

Figure 4: Research Count Based on Object Predicted

The research objects were classified into five, namely Index, Stock, Commodity, Currency, and ETF as show in Figure 4. Stock index is an indicator that describes market conditions at a time. For instance, when stock price index experiences an increase, a majority of stock prices within it also tend to rise. Several research used stock index as the object of investigation [62][29][63][64]. Research using stock price index as an object is still widely investigated by 41.3% of researchers.

The share price is the price established by a firm for those who seek to acquire share ownership rights. The valuation of stock prices fluctuates continuously. The supply and demand between sellers and buyers of shares influence the share price value. Numerous research using stock prices as an object is shown in [65][66][21][67][23][26].

Commodities tangible goods with easy tradability, the potential for physical distribution, and exchangeability for similar products are commonly exchanged through futures exchanges. Typically, investors can purchase or sell commodities through futures exchanges. Several investigations have taken topics with commodity objects, as researched by Kanwal et al. [68]. However, this research simultaneously examines the DAX index and the Hang Seng Index as research objects.

The exchange rate is the quantity of local currency required to purchase one unit of different currencies. Tian et al. conducted stock price prediction using the LSTM and Light-GBM hybrid models [36]. The objective of this research is to forecast stock prices, indexes, and USDJPY currency fluctuations.

Exchange Traded Funds (ETFs) are mutual funds whose performance mirrors specific indexes and can be traded like stock on exchanges. Lee et al. explored the efficacy of deep neural networks (DNN) with the technical indicator implemented to stock forecasts in predicting the TWSE 0050 market's direction [69].

3.3 Dataset Characteristics

3.3.1 Covered Stock Exchange and Country

This research investigates a variety of worldwide indices and stock prices, directly or indirectly collecting input data [70] and subsequently employing this data to forecast future values. Some research focuses on forecasting stock indices at specific moments to determine the associated risks, which raises concerns about the dependability of market resilience for highly profitable investments. Such circumstances can significantly contribute to the accuracy of forecasting methods, stock market, and stock market returns [71].

Table 4 presents an overview of the market under scrutiny along with their corresponding major stock market. While most investigations focus on a single market, some evaluate their methods on many market. In the past five years, established market (e.g., the United States) and emerging market (e.g., China and India) have attracted significant attention from the academic community.

Table 4: The Researched Market And The Respective Major Index Of Stock Market.

Country	Stock Exchange	Stock Index	Article Count
England	London Stock Exchange	FTSE	3
India	National Stock	NIFTY	14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

	Exchange		
	Bombay Stock Exchange	BSE SENSEX	8
US	New York Stock Exchange	S&P 500, DJIA	38
	Nasdaq Stock Exchange	NASDAQ	17
Canada	Toronto Stock Exchange	S&P/TSX	1
China	Shanghai Stock Exchange	SSEC	29
	Shenzhen Stock Exchange	SZSE Component	9
Hong Kong	Hong Kong Stock Exchange	HSI	7
Korea	Korea Exchange	KOSPI, KOSDAQ	6
Japan	Tokyo Stock Exchange	Nikkei, TOPIX	4
Taiwan	Taiwan Stock Exchange	TAIEX	3
Indonesia	Indonesistock Exchange	IDX Composite	1
German	Frankfurt Stock Exchange	DAX	4
Turkey	Borsa Istanbul Stock Exchange	BIST	2
Pakistan	Pakistan Stock Exchange	KSE, KMI	1
France	Euronext Paris Stock Exchange	CAC	1
Iraq	Iraq Stock Exchange	ISX	1

Table 4 and Figure 5 shows the predominant focus of research on the New York Stock Exchange, followed by the Shanghai Stock Exchange in second place, and the Nasdaq Stock Exchange in third place. When considering the surveyed nations as shown in Figure 6, the United States, China, and India consistently maintain their top positions as the most extensively researched stock exchange environments. Developing countries and emerging market such as Indonesia, Pakistan, and Iraq still receive attention from researchers at 0.67%. Maqsood et al. conducted research to predict indices in the US, Hong Kong, Turkey, and Pakistan using Deep Learning [72]. Diqi, Hiswati, and Nur use the GAN algorithm to predict Aneka Tambang's shares traded on the Indonesistock Exchange [38]. Abdulhussein et al. conducted research to predict the Iraqi Stock Market using the Convolution Neural Network [53].

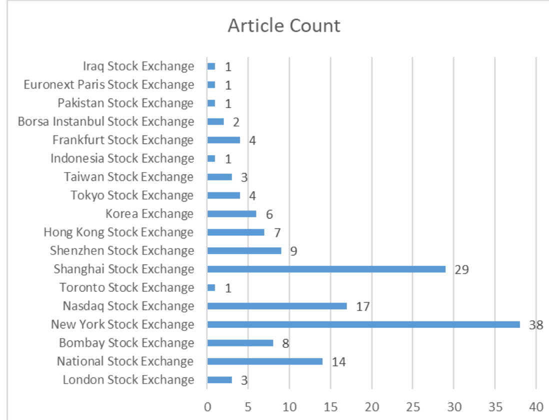

Figure 5: The Amount of Research Based on The Stock Market

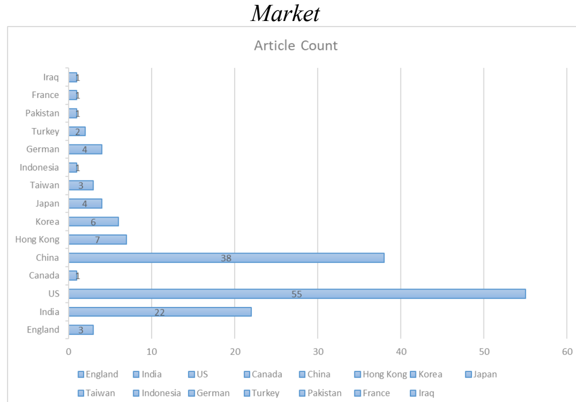

Figure 6: The Amount of Research Based on The Country Where The Stock Exchange is Located

3.3.2 Data Length

The evaluation of the effectiveness of different models requires historical data. However, there is a trade-off when determining the data length. An inadequate sample size proves insufficient in validating efficacy, bearing a heightened risk of overfitting. Conversely, an extended timeframe introduces the hazard of encompassing diverse market conditions, potentially leading to conclusions that lack current relevance. Data availability and affordability must be considered when choosing the data length. Good quality intraday data is more expensive, and most past research, including intraday prediction, used a time range of less than one year.

This review shows that the percentage of data lengths 1-5 years, 5-10 years, and 10-25 are almost the same. Conversely, data lengths of less than 1 year and over 25 years are infrequently used by researchers. Research [73] uses the most extended dataset, about 92 years for the S&P500 index, 124 years for the DJIA index, and about 50 years for the NASDAQ index. Yin et al. used a 3-month dataset to predict China A-share on a high-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

frequency basis [74]. Figure 7 shows the number of studies based on the dataset length.

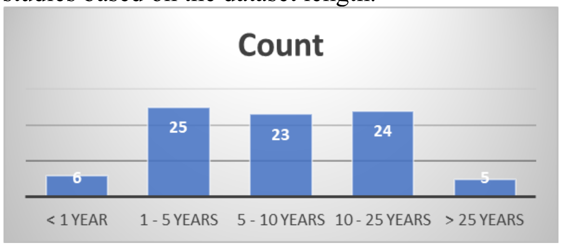

Figure 7: The Number of Studies Based on The Dataset Length

3.3.3 Data Frequency

The frequency at which data is collected plays a pivotal role in determining the volume of data amassed within a specific timeframe. The data items used for stock prediction can be represented in several ways, from historical data (open, high, low, and close) to relevant metrics. The frequency of data can range from a fraction of a second to minutes, hours, daily, or even yearly. Second-frequency data is often associated with prediction process at high frequencies. A more typical presentation of stock market data as bars refers to the presentation of several sample points as an intelligible aggregation of the period's highlights. Other data sources commonly used are non-traditional data sources. This news source can be in the form of text data, namely news and social media. This data source is used for stock prediction in the same time frame. Beyond conventional sources, non-traditional data sources, including textual data such as news articles and social media posts, have gained traction as viable resources. These sources often harnessed for stock prediction in real-time, have emerged as input variables in prediction systems. For instance, Day and Lee illustrate this by using an array of daily media headlines as training data [75].

For data frequency, the stock/index data frequency was classified into seconds, minutes, hours, daily, weekly, monthly, and yearly. Based on the papers reviewed, it can be concluded that many researchers conducted research using daily data. Few researchers use data in seconds and minutes frequency. Research that uses annual data frequency is also carried out sparingly because it is difficult to obtain data. The use of annual data will produce prediction for a very long period. Figure 8 shows the distribution of the number of studies based on dataset frequency.

Research conducted by Yin et al. uses data taken every 3 seconds [74]. This research develops the Graph Attention LSTM algorithm, which forecast trends in the Chinstock market, specifically focusing on high-frequency characteristics. Similarly, Long, Lu, and Cui use data with a 1-minute frequency to forecast the direction of the Chinese stock market index (CSI 300) [76]. Some researchers predict stock market by using data in a 5-minute frequency [77][78]. Most research is carried out using daily data, which is 84.1%, as exemplified by previous works by [79][20][80][81].

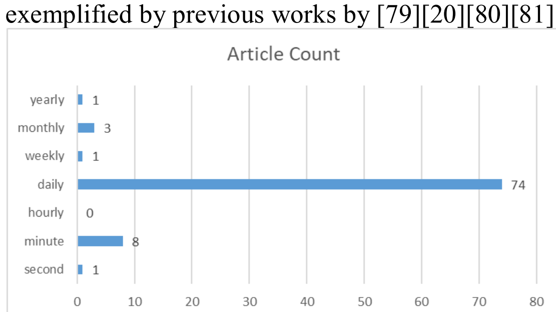

Figure 8: The Number of Studies Based on Dataset Frequency

3.4 Prediction Methods

The majority of predictive models employ supervised learning as their foundation. In supervised learning, a training dataset is used for learning, while a separate testing dataset is employed for evaluation purposes. Predictive models can be categorized into traditional models and alternative models. Among the traditional models, commonly employed ones include feedforward neural networks, CNN, and recurrent neural networks (RNN). This section will focus on its application in the field of stock prediction and not focus on the theory of Deep Learning algorithms.

3.4.1 Feedforward Neural Network

The most basic form of an artificial neural network is one in which nodes are not interconnected cyclically. An ANN neuron consists of a function that counts the total input and an activation function that provides outputs. A subset of ANN comprising both input and output layers with the same number of nodes is an Auto-Encoder (AE). In this literature review, ANNs with two or more hidden layers are called DNN. Backpropagation Neural Networks (BPNN), Multilayer Perceptron (MLP), and Extreme Learning Machines (ELM) are examples of Feedforward Neural Networks. This research discovered a variety of FFNN-based investigations, which are included in Table 5.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5: Research That Use FFNN.

Algorithm	Research
ANN	[82]
DNN	[57][65][35][77]
BPNN	[24][46]
EML	[25][83]

3.4.2 CNN

A CNN is a type of ANN that is commonly used for data analysis, classification, and recognition. This network structure comprises multiple layers of processing units, organized hierarchically, with each layer progressively acquiring more intricate features from the input data. The key feature of CNNs is the convolution operation, which is a mathematical operation that allows the network to learn features from raw data. Research using the CNN basis is found in the paper [26][84][85][53][86].

3.4.3 RNN

A RNN is a type of neural network architecture designed for processing sequences of data by maintaining a hidden state that evolves, enabling the network to capture temporal dependencies and patterns in the data. The LSTM network is a RNN that overcomes the problem of vanishing gradients by replacing hidden layers with forget gates. Gated Recurrent Units (GRU) is an additional RNN that employs a forget gate but has fewer parameters than LSTM. A bidirectional RNN, as the name suggests, combines two hidden layers that process data in opposite directions, ultimately producing a singular output. Both two-way LSTM (Bi-LSTM) and two-way GRU (BGRU) have been used for stock forecasting. In Table 6, this research identifies various publications employing the RNN method.

Table 6: Research That Use RNN.

Algorithm	Research
LSTM	[48][59][79][81][22][39][87][47][88][55][89][69][90]
Bi-LSTM	[28][91][92]
GRU	[93]

3.4.4 Ensemble Algorithm

Ensemble learning is a method in which several models are trained using several algorithms and combined. The goal of ensemble learning is to increase efficiency. Although classical models have long been used, deep learning models dominate with superior performance. Combining these two approaches' advantages, ensemble deep learning can provide superior performance [94]. Over time, ensemble techniques have evolved to strengthen the generalization of the models created. Figure 9 illustrates the various categories of ensemble strategies that have been developed.

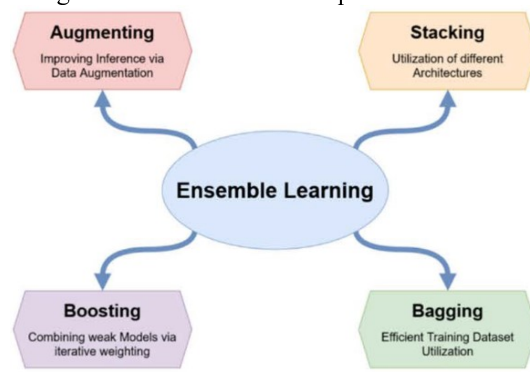

Figure 9: Ensemble Model [95]

Li and Pan [61] introduce a strategy based on Deep Learning to forecast future stock price movements. This research used a mixed ensemble learning method that combines two RNN and a completely connected neural network, and the testing phases were carried out using S&P 500 index. Experiments show that integrating ensemble Deep Learning models can outperform existing best prediction strategies. This method reduces mean-squared error, increases precision rate, increases recall by 50%, and F1-score by 44.78%. This research shows that Deep Learning ensemble systems that have been researched can predict future market trends more accurately and assist investors in making better financial decisions than conventional methods.

Zhao and Cheng [96], employing an ensemble learning methodology, refine and amalgamate diverse linear and nonlinear individual stock return prediction models. In US market forecasting applications, stacking with a simple structure outperforms traditional historical average benchmarks. In addition, this analysis shows that out-of-sample gains from piling are most pronounced when market experiences large negative moves. Collectively, the use of compounding can notably heighten the predictability of excess market return.

Xiao and Su use classical and machine learning methods to anticipate linear and nonlinear issues, respectively, to forecast stock time-series data. First, between 2010 and 2019, stock data are gathered from the US market. Then, the ARIMA and LSTM algorithms are used for training and forecasting stock price and its sub-correlation. This research obtained the results: (1) ARIMA and LSTM algorithms can accurately predict stock prices and their correlations; (2) In stock price prediction, the LSTM algorithm can outperform the ARIMA model; and (3) the ARIMA-LSTM ensemble model outperforms other benchmark methods by a substantial margin. Consequently, this recommended strategy helps investors with the theoretical basis and method recommendations for stock trading in the Chinstock exchange [97].

3.4.5 Hybrid Models

3.4.5.1 *CNN-LSTM* [98]

This research includes two lines, one built on CNN modules and the other including CNN and Bi-LSTM lines. These pathways converge at a multilayer fusion center, which integrates local features, serving as the center of fusion. This research used the PRIMO COVID-19 dataset, which contains data on the Dow Jones stock index and related tweets from January 1, 2016, to July 30, 2020. In addition to evaluating precision, this research includes sensitivity and specificity metrics to evaluate several error types. Subsequently, sensitivity is the number of occasions when the model successfully forecasts an increase in market prediction results. To present the performance advantages of the framework proposed, a performance assessment was carried out using the PRIMO COVID-19 dataset.

3.4.5.2 *Fast RNN - CNN – Bi-LSTM* [99]

This research proposes two different models, each model is used for a different purpose. The first method uses Fast Recurrent Neural Networks (Fast RNNs). This method was first applied to stock price forecasting. The next method is a hybrid Deep Learning method that integrates the great attributes of Fast-RNNs, CNNs, and Bi-LSTM models to predict short stock market volatility. The investigation concentrates on 1-day and 3-day stock data from four companies, sampled at 1-minute intervals. This method shows superior performance, boasting a reduced Root Mean Square Error (RMSE). Additionally, its computational efficiency positions it for real-time applications, offering practical recommendations. This model outperforms the ARIMA, LSTM, and other proposed hybrid models for direct stock price forecasting.

3.4.5.3 *CNN-Bi-LSTM* [100]

In this research, a novel CNN-Bi-SLSTM model is introduced to forecast the closing prices of stock. The Bidirectional Special LSTM (Bi-SLSTM) forms an enhancement over the Bi-LSTM model, incorporating a hyperbolic tangent function in the output unit to bolster predictive capabilities concerning stock values. The method gathers advanced components that affect stock price using a CNN and forecast daily stock value using Bi-SLSTM after CNN processing of the data. The CNN-Bi-SLSTM is trained and assessed using Shenzhen Stock Exchange Index data from July 1991 to October 2020, to establish the model's effectiveness. This research performs comparative analyses of CNN-Bi-SLSTM against four established methodologies. In the first comparison, the FFNN method, namely MLP, will be used. Second, CNN-Bi-SLSTM will be compared with the RNN and LSTM methods. Finally, CNN-Bi-SLSTM will also be compared with the CNN-LSTM hybrid method. As indicated by the experimental data, the MAE, RMSE, and R-square (R2) are all perfect. Therefore, CNN-Bi-SLSTM shows remarkable accuracy in forecasting the SZSE Index's daily prices one day in advance, offering traders a tool to mitigate risks effectively.

3.4.5.4 *Deep CNN – reinforcement LSTM* [80]

For forecasting future stock prices, deep CNN and a reinforcement-LSTM model are created using vast quantities of data. Furthermore, real-time stock future prices from the US, London, Taiwan, and Indistock market are harnessed to comprehensively appraise the model's efficacy. This research is evaluated by predicting stock prices one month, one week, and one day ahead. A whole year's worth of data is gathered, and tests based on the proposed model are carried out. The modeling results reveal that the proposed model outperforms current methods in terms of many parameters, including POCID values greater than 85 percent, R2 values greater than 80 percent, ARV values less than 0.024 percent, and MAPE values less than 0.010 percent.

3.4.5.5 *DNN – LSTM – 1D CNN* [68]

This research introduces a hybrid Deep Learning (DL) system that combines a Bidirectional Cuda Deep Neural Network LSTM (Bi-Cu-DNN-LSTM) with a 1D CNN for accurate and speedy stock price forecasting. For validation, this proposed model (Bi-Cu-DNN-LSTM-1dCNN) was compared with conventional and hybrid Deep Learning models using five stock price data sets. The projected results indicate that this model tested is sufficient for precise stock price forecasting and supporting investors in making informed investment decisions.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.4.6 Other Model

Other models include Generative Adversarial Networks (GAN), transfer learning, and reinforcement learning. Until now, this GAN model has not been widely used for stock forecasting. The use of this model is still in the early stages and the development stage. Several research efforts involving GANs can be identified, specifically documented in works by [34][38][101].

A comparative summary of stock prediction research using Deep Learning can be seen in Table 7 at the appendix.

4. CONCLUSION AND FUTURE WORK

In reflecting upon our work, it's imperative to acknowledge its scope and inherent limitations. While our research offers an encompassing overview of methodologies and trends in stock prediction using deep learning from 2017 to 2022, it's confined by the temporal scope, potentially omitting foundational works or significant breakthroughs preceding this period. Our reliance on specific academic journals and conferences as data sources may inadvertently neglect pertinent studies published elsewhere. The rigorous criteria employed in our study, although methodical, are not immune to subjectivity, which could introduce nuances of bias in article selection and interpretation. Furthermore, given the rapidly evolving nature of the field of stock prediction, our findings, though current, may soon be succeeded by newer methodologies and trends. A notable caveat is the absence of practical implementation or validation within our study, confining our research to a literature-based purview and necessitating empirical validation for real-world applicability. Recognizing these constraints not only underscores the depth of our introspection but also paves the way for more exhaustive, unbiased, and updated research in the future.

Our study has meticulously analyzed 86 articles from the 2017-2022 period, offering a comprehensive review of methodologies and trends in stock prediction using deep learning. While it captures the prevailing practices and insights, it's essential to note the dynamic nature of artificial intelligence and finance. Rapid advancements could soon introduce new paradigms, highlighting the transient nature of our current landscape snapshot.

The inherent nature of our literature review, being secondary research, leans on existing knowledge without generating new empirical data. This means our conclusions are reliant on the quality and accuracy of the studies reviewed. Any biases or oversights in those papers could inadvertently influence our findings. This underscores the continuous need for empirical research in the realm of stock prediction using deep learning, and the importance of validation and experimentation in the field.

To summarize, this study endeavored to elucidate the input features, output characteristics, attributes of datasets, methodologies, and trends associated with hybrid and ensemble models, in addition to the evaluation methods employed in stock prediction through Deep Learning from 2017–2022. By implementing strict exclusion criteria, we were able to identify 86 pertinent articles from the aforementioned period. This work, underpinned by an exhaustive Literature Review, ensures a meticulous and comprehensive exposition. The Literature Review, in this context, serves as a structured approach to discern, assimilate, and interpret variables critical to addressing the research inquiries.

During the analysis of the input features used, the following were obtained, 53.66% of research uses only historical trading data as a system input. 24.39% of researchers focused on Historical Trading and Technical Indicator Data as input. 3.66% of articles used a combination of Historical Trading and Fundamental Data as input parameters. 4.88% of researchers focused on input using a combination of Historical Trading, Technical Indicators, and Text Data. 7.32% of publications used a combination of Historical Trading and Text Data. Only 2.44% each used text data or a combination of Historical Trading and Macroeconomics. Lastly, only 1.22% of researchers used a combination of Historical Trading, Fundamental, and Text Data as system input parameters.

Regarding the output variables used, three research topics were classified: market trends, price prediction, and both. A substantial proportion of 58.82% of researchers focused their investigations on predicting market trends. Meanwhile, only 7.06% of the research conducted both research predicting market trends and predicting stock prices. Based on the objects researched, five objects were identified, namely stock, indexes, commodities, currencies, and ETFs. Although literature review only covers stock and index prediction, commodity, currency, and ETF predictions are often used as benchmark data during testing. The percentage of research using stock, indices, commodities, currencies, and ETF objects

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

was 54.35%, 41.30%, 1.09%, 2.18%, and 1.09%, respectively.

Based on the dataset's characteristics, the analysis aimed to examine the distribution of stock market and the countries researched, the distribution of data intervals, and the distribution of data frequencies. Examining the distribution of stock exchange data, it is evident that the New York Stock Exchange and the Shanghai Stock Exchange continue to be the primary focal points for researchers. The US, China, and India are still the three countries whose exchanges are very active and attract the attention of many researchers. The US represents countries with mature economies, while China and India represent emerging market. Even though the capitalization is not large, several countries in Asia are also of concern in several research, namely Taiwan, Indonesia, Iraq, and Pakistan. Based on the length of the data, it was found that data lengths of 1-5 years, 5-10 years, and 10-25 years have an almost uniform distribution, which is around 30%. While for data length > 1 year, it is only 7.2%, and for data > 25 years, it is only 6.02%. Data with a length of over 25 years is almost certain to be challenging to obtain. Based on the frequency of data, it was found that almost all researchers use daily data as a basis for the frequency of data used. The use of seconds or minutes of data was used for trading in the concise term. In contrast, weekly, monthly, and even annual data are used by investors with very long-time intervals.

Based on prediction methods, these stock prediction methods are grouped into FFNN, CNN, RNN, hybrid, ensemble, and other (GAN, transfer learning, and reinforcement learning) methods. Algorithms based on LSTM and Bi-LSTM are still excellent among researchers because of their ability to remember time series. Algorithms that showed potential for further development included hybrid and ensemble-based methods.

Based on the observed characteristics, exploration of input combinations involving multiple parameters remains infrequent, particularly those integrating text and historical trading data. There are still opportunities for improving the accuracy of stock forecasts using a combination of historical data input, fundamental data, and text data comparable to that used by [47].

Considering the investigations carried out in the country's stock market that have garnered less attention. There are also chances to perform research using high-frequency data (seconds, minutes, hours), as shown by [74] [76] [77] [78]. One of the challenges lies in selecting an accurate and efficient algorithm.

Given the predictive methodology's potency, there's potential for optimizing extensively employed methods in financial market prediction. Methods such as Bayesian [36], genetic algorithms [22], and wavelet transform [62] are among the methods that have been investigated. There are also opportunities for combining many state-of-the-art algorithms to create ensemble or hybrid algorithms.

REFERENCES:

[1] E. Sudarmanto, *Pasar Uang dan Pasar Modal*, vol. 1, no. September. Penerbit Yayasan Kita Menulis, 2021.

[2] B. D. Susilo, *Pasar Modal, Mekanisme Perdagangan Saham, Analisis Sekuritas, dan Strategi Investasi Di Bursa Efek Indonesia BEI*. 2009.

[3] K. Levi, "Transfer of Innovation „ Development and Approbation of Applied Courses Based on the Transfer of Teaching Innovations in Finance and Management for Further Education of Entrepreneurs and Specialists in Latvia , Lithuania and Bulgaria "," *Educ. Cult. Lifelong Learn. Program.*, pp. 1–166, 2010, [Online]. Available: http://www.bcci.bg/projects/latvia/pdf/8_IAPM_final.pdf

[4] M. A. Lim, *The Handbook of Technical Analysis*. John Wiley & Sons Singapore Pte. Ltd, 2016.

[5] N. Naik and B. R. Mohan, "Intraday Stock Prediction Based on Deep Neural Network," *Natl. Acad. Sci. Lett.*, vol. 43, no. 3, pp. 241–246, 2020, doi: 10.1007/s40009-019-00859-1.

[6] X. Cui, W. Shang, F. Jiang, and S. Wang, "Stock Index Forecasting by Hidden Markov Models with Trends Recognition," *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019*, pp. 5292–5297, 2019, doi: 10.1109/BigData47090.2019.9006068.

[7] W. Jiang, "Applications of deep learning in stock market prediction: Recent progress," *Expert Syst. Appl.*, vol. 184, no. July, p. 115537, 2021, doi: 10.1016/j.eswa.2021.115537.

[8] M. Velay and F. Daniel, "Stock Chart Pattern recognition with Deep Learning," 2018.

[9] Z. Hu, Y. Zhao, and M. Khushi, "A survey of forex and stock price prediction using deep learning," *Appl. Syst. Innov.*, vol. 4, no. 1, pp.

1–30, 2021, doi: 10.3390/ASI4010009.

[10] C. Zhang, N. N. A. Sjarif, and R. B. Ibrahim, "Decision Fusion for Stock Market Prediction: A Systematic Review," *IEEE Access*, vol. 10, no. July, pp. 81364–81379, 2022, doi: 10.1109/ACCESS.2022.3195942.

[11] J. Shah, D. Vaidya, and M. Shah, "A comprehensive review on multiple hybrid deep learning approaches for stock prediction," *Intell. Syst. with Appl.*, vol. 16, no. September 2021, p. 200111, 2022, doi: 10.1016/j.iswa.2022.200111.

[12] S. Maddodi and K. G. N. Kumar, "Stock Market Forecasting: a Review of Literature," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. Vol 5, No, p. 11, 2021, [Online]. Available: https://ojs.stmikpringsewu.ac.id/index.php/ijiscs/article/download/1064/pdf

[13] G. Y. Dopi, R. Hartanto, and S. Fauziati, "Systematic Literature Review: Stock Price Prediction Using Machine Learning and Deep Learning," in *2021 IEEE Mysore Sub Section International Conference (MysuruCon)*, Oct. 2021, vol. 194, no. Icombest, pp. 660–664. doi: 10.2991/aebmr.k.211117.008.

[14] B. Kitchenham, "Procedures for Performing Systematic Reviews, Version 1.0," *Empir. Softw. Eng.*, vol. 33, no. 2004, pp. 1–26, 2004, [Online]. Available: https://www.researchgate.net/profile/Barbara-Kitchenham/publication/228756057_Procedures_for_Performing_Systematic_Reviews/links/618cfae961f09877207f8471/Procedures-for-Performing-Systematic-Reviews.pdf

[15] B. Kitchenham and S. M. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering Guidelines for performing Systematic Literature Reviews in Software Engineering EBSE Technical Report EBSE-2007-01 Software Engineering Group School of Computer Science and Ma," no. October 2021, 2007.

[16] G. Spanos and L. Angelis, "The impact of information security events to the stock market: A systematic literature review," *Comput. Secur.*, vol. 58, pp. 216–229, 2016, doi: 10.1016/j.cose.2015.12.006.

[17] K. Kumar and M. T. U. Haider, *Enhanced Prediction of Intra-day Stock Market Using Metaheuristic Optimization on RNN–LSTM Network*, vol. 39, no. 1. Ohmsha, 2021. doi: 10.1007/s00354-020-00104-0.

[18] A. Thakkar, D. Patel, and P. Shah, "Pearson Correlation Coefficient-based performance enhancement of Vanilla Neural Network for stock trend prediction," *Neural Comput. Appl.*, vol. 33, no. 24, pp. 16985–17000, 2021, doi: 10.1007/s00521-021-06290-2.

[19] Y. Li, P. Ni, and V. Chang, "Application of deep reinforcement learning in stock trading strategies and stock forecasting," *Computing*, vol. 102, no. 6, pp. 1305–1322, 2020, doi: 10.1007/s00607-019-00773-w.

[20] S. Bhanja and A. Das, "A Black Swan event-based hybrid model for Indian stock markets' trends prediction," *Innov. Syst. Softw. Eng.*, 2022, doi: 10.1007/s11334-021-00428-0.

[21] W. Chen, H. Zhang, M. K. Mehlawat, and L. Jia, "Mean–variance portfolio optimization using machine learning-based stock price prediction," *Appl. Soft Comput.*, vol. 100, p. 106943, 2021, doi: 10.1016/j.asoc.2020.106943.

[22] A. Thakkar and K. Chaudhari, "Information fusion-based genetic algorithm with long short-term memory for stock price and trend prediction," *Appl. Soft Comput.*, vol. 128, p. 109428, 2022, doi: 10.1016/j.asoc.2022.109428.

[23] N. Jing, Z. Wu, and H. Wang, "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction," *Expert Syst. Appl.*, vol. 178, no. March, p. 115019, 2021, doi: 10.1016/j.eswa.2021.115019.

[24] X. Liu, J. Guo, H. Wang, and F. Zhang, "Prediction of stock market index based on ISSA-BP neural network," *Expert Syst. Appl.*, vol. 204, no. May, p. 117604, 2022, doi: 10.1016/j.eswa.2022.117604.

[25] D. Wu, X. Wang, and S. Wu, "A hybrid framework based on extreme learning machine, discrete wavelet transform, and autoencoder with feature penalty for stock prediction," *Expert Syst. Appl.*, vol. 207, no. July 2021, p. 118006, 2022, doi: 10.1016/j.eswa.2022.118006.

[26] W. Chen, M. Jiang, W. G. Zhang, and Z. Chen, "A novel graph convolutional feature based convolutional neural network for stock trend prediction," *Inf. Sci. (Ny).*, vol. 556, pp. 67–94, 2021, doi: 10.1016/j.ins.2020.12.068.

[27] A. F. Kamara, E. Chen, and Z. Pan, "An ensemble of a boosted hybrid of deep learning models and technical analysis for forecasting

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

stock prices," *Inf. Sci. (Ny).*, vol. 594, pp. 1–19, 2022, doi: 10.1016/j.ins.2022.02.015.

[28] X. Chen, X. Ma, H. Wang, X. Li, and C. Zhang, "A hierarchical attention network for stock prediction based on attentive multi-view news learning," *Neurocomputing*, vol. 504, pp. 1–15, 2022, doi: 10.1016/j.neucom.2022.06.106.

[29] 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," *Phys. A Stat. Mech. its Appl.*, vol. 541, no. 258, p. 122272, 2020, doi: 10.1016/j.physa.2019.122272.

[30] S. Wu, Y. Liu, Z. Zou, and T. H. Weng, "S_I_LSTM: stock price prediction based on multiple data sources and sentiment analysis," *Conn. Sci.*, vol. 34, no. 1, pp. 44–62, 2022, doi: 10.1080/09540091.2021.1940101.

[31] B. L. Shilpa and B. R. Shambhavi, "Combined deep learning classifiers for stock market prediction: integrating stock price and news sentiments," *Kybernetes*, 2021, doi: 10.1108/K-06-2021-0457.

[32] S. M. Prabin and M. S. Thanabal, "A repairing artificial neural network model-based stock price prediction," *Int. J. Comput. Intell. Syst.*, vol. 14, no. 1, pp. 1337–1355, 2021, doi: 10.2991/IJCIS.D.210409.002.

[33] M. Sharaf, E. E. D. Hemdan, A. El-Sayed, and N. A. El-Bahnasawy, "StockPred: a framework for stock Price prediction," *Multimed. Tools Appl.*, vol. 80, no. 12, pp. 17923–17954, 2021, doi: 10.1007/s11042-021-10579-8.

[34] A. Kumar *et al.*, "Generative adversarial network (GAN) and enhanced root mean square error (ERMSE): deep learning for stock price movement prediction," *Multimed. Tools Appl.*, vol. 81, no. 3, pp. 3995–4013, 2022, doi: 10.1007/s11042-021-11670-w.

[35] S. Il Lee and S. J. Yoo, "Multimodal deep learning for finance: integrating and forecasting international stock markets," *J. Supercomput.*, vol. 76, no. 10, pp. 8294–8312, 2020, doi: 10.1007/s11227-019-03101-3.

[36] L. Tian, L. Feng, L. Yang, and Y. Guo, "Stock price prediction based on LSTM and LightGBM hybrid model," *J. Supercomput.*, vol. 78, no. 9, pp. 11768–11793, 2022, doi: 10.1007/s11227-022-04326-5.

[37] M. Li, Y. Zhu, Y. Shen, and M. Angelova, "Clustering-enhanced stock price prediction using deep learning," *World Wide Web*, 2022, doi: 10.1007/s11280-021-01003-0.

[38] M. Diqi, M. E. Hiswati, and A. S. Nur, "StockGAN: robust stock price prediction using GAN algorithm," *Int. J. Inf. Technol.*, vol. 14, no. 5, pp. 2309–2315, 2022, doi: 10.1007/s41870-022-00929-6.

[39] A. M. Rather, "LSTM-based Deep Learning Model for Stock Prediction and Predictive Optimization Model," *EURO J. Decis. Process.*, vol. 9, no. September, p. 100001, 2021, doi: 10.1016/j.ejdp.2021.100001.

[40] S. Carta, A. Ferreira, A. S. Podda, D. Reforgiato Recupero, and A. Sanna, "Multi-DQN: An ensemble of Deep Q-learning agents for stock market forecasting," *Expert Syst. Appl.*, vol. 164, no. November 2019, p. 113820, 2021, doi: 10.1016/j.eswa.2020.113820.

[41] H. Rezaei, H. Faaljou, and G. Mansourfar, "Stock price prediction using deep learning and frequency decomposition," *Expert Syst. Appl.*, vol. 169, no. October 2020, p. 114332, 2021, doi: 10.1016/j.eswa.2020.114332.

[42] Y. Chen, J. Wu, and Z. Wu, "China's commercial bank stock price prediction using a novel K-means-LSTM hybrid approach," *Expert Syst. Appl.*, vol. 202, no. August 2021, p. 117370, 2022, doi: 10.1016/j.eswa.2022.117370.

[43] G. Liu and W. Ma, "A quantum artificial neural network for stock closing price prediction," *Inf. Sci. (Ny).*, vol. 598, pp. 75–85, 2022, doi: 10.1016/j.ins.2022.03.064.

[44] M. Tao, S. Gao, D. Mao, and H. Huang, "Knowledge graph and deep learning combined with a stock price prediction network focusing on related stocks and mutation points," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 7, pp. 4322–4334, 2022, doi: 10.1016/j.jksuci.2022.05.014.

[45] Q. Liu, Z. Tao, Y. Tse, and C. Wang, "Stock market prediction with deep learning: The case of China," *Financ. Res. Lett.*, vol. 46, no. PA, p. 102209, 2022, doi: 10.1016/j.frl.2021.102209.

[46] D. Zhang and S. Lou, "The application research of neural network and BP algorithm in stock price pattern classification and prediction," *Futur. Gener. Comput. Syst.*, vol. 115, pp. 872–879, 2021, doi: 10.1016/j.future.2020.10.009.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[47] Q. Li, J. Tan, J. Wang, and H. Chen, "A Multimodal Event-Driven LSTM Model for Stock Prediction Using Online News," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 10, pp. 3323–3337, 2021, doi: 10.1109/TKDE.2020.2968894.

[48] T. Swathi, N. Kasiviswanath, and A. A. Rao, "An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis," *Appl. Intell.*, pp. 13675–13688, 2022, doi: 10.1007/s10489-022-03175-2.

[49] Q. Q. He, S. W. I. Siu, and Y. W. Si, "Instance-based deep transfer learning with attention for stock movement prediction," *Appl. Intell.*, 2022, doi: 10.1007/s10489-022-03755-2.

[50] W. Khan, M. A. Ghazanfar, M. A. Azam, A. Karami, K. H. Alyoubi, and A. S. Alfakeeh, "Stock market prediction using machine learning classifiers and social media, news," *J. Ambient Intell. Humaniz. Comput.*, vol. 13, no. 7, pp. 3433–3456, 2022, doi: 10.1007/s12652-020-01839-w.

[51] D. Chandola, A. Mehta, S. Singh, V. A. Tikkiwal, and H. Agrawal, "Forecasting Directional Movement of Stock Prices using Deep Learning," *Ann. Data Sci.*, 2022, doi: 10.1007/s40745-022-00432-6.

[52] S. P. Chatzis, V. Siakoulis, A. Petropoulos, E. Stavroulakis, and N. Vlachogiannakis, "Forecasting stock market crisis events using deep and statistical machine learning techniques," *Expert Syst. Appl.*, vol. 112, pp. 353–371, 2018, doi: 10.1016/j.eswa.2018.06.032.

[53] S. H. Abdulhussein, N. J. Al-Anber, and H. A. Atee, "Iraqi Stock Market Prediction using Proposed Model of Convolution Neural Network," *J. Comput. Sci.*, vol. 18, no. 5, pp. 350–358, 2022, doi: 10.3844/jcssp.2022.350.358.

[54] J. Zhang and Y. Lei, "Deep Reinforcement Learning for Stock Prediction," *Sci. Program.*, vol. 2022, no. Ml, 2022, doi: 10.1155/2022/5812546.

[55] M. Iyyappan, S. Ahmad, S. Jha, A. Alam, M. Yaseen, and H. A. M. Abdeljaber, "A Novel AI-Based Stock Market Prediction Using Machine Learning Algorithm," *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/4808088.

[56] M. Jiang, L. Jia, Z. Chen, and W. Chen, "The two-stage machine learning ensemble models for stock price prediction by combining mode decomposition, extreme learning machine and improved harmony search algorithm," *Ann. Oper. Res.*, vol. 309, no. 2, pp. 553–585, 2022, doi: 10.1007/s10479-020-03690-w.

[57] Y. Song, J. W. Lee, and J. Lee, "A study on novel filtering and relationship between input-features and target-vectors in a deep learning model for stock price prediction," *Appl. Intell.*, vol. 49, no. 3, pp. 897–911, 2019, doi: 10.1007/s10489-018-1308-x.

[58] J. Ni and Y. Xu, "Forecasting the Dynamic Correlation of Stock Indices Based on Deep Learning Method," *Comput. Econ.*, 2021, doi: 10.1007/s10614-021-10198-3.

[59] C. Ma, Y. Liang, S. Wang, and S. Lu, "Stock linkage prediction based on optimized LSTM model," *Multimed. Tools Appl.*, vol. 81, no. 9, pp. 12599–12617, 2022, doi: 10.1007/s11042-022-12381-6.

[60] J. Wang, Q. Cheng, and Y. Dong, "An XGBoost-based multivariate deep learning framework for stock index futures price forecasting," *Kybernetes*, 2022, doi: 10.1108/K-12-2021-1289.

[61] Y. Li and Y. Pan, "A novel ensemble deep learning model for stock prediction based on stock prices and news," *Int. J. Data Sci. Anal.*, vol. 13, no. 2, pp. 139–149, 2022, doi: 10.1007/s41060-021-00279-9.

[62] M. Zolfaghari and S. Gholami, "A hybrid approach of adaptive wavelet transform, long short-term memory and ARIMA-GARCH family models for the stock index prediction," *Expert Syst. Appl.*, vol. 182, no. November 2020, p. 115149, 2021, doi: 10.1016/j.eswa.2021.115149.

[63] S. Mukherjee, B. Sadhukhan, N. Sarkar, D. Roy, and S. De, "Stock market prediction using deep learning algorithms," *CAAI Trans. Intell. Technol.*, no. February, 2021, doi: 10.1049/cit2.12059.

[64] Z. Fathali, Z. Kodia, and L. Ben Said, "Stock Market Prediction of NIFTY 50 Index Applying Machine Learning Techniques," *Appl. Artif. Intell.*, vol. 36, no. 1, 2022, doi: 10.1080/08839514.2022.2111134.

[65] R. Singh and S. Srivastava, "Stock prediction using deep learning," *Multimed. Tools Appl.*, vol. 76, no. 18, pp. 18569–18584, Sep. 2017, doi: 10.1007/s11042-016-4159-7.

[66] J. Long, Z. Chen, W. He, T. Wu, and J. Ren, "An integrated framework of deep learning and knowledge graph for prediction of stock

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

price trend: An application in Chinese stock exchange market," *Appl. Soft Comput. J.*, vol. 91, p. 106205, 2020, doi: 10.1016/j.asoc.2020.106205.

[67] A. U. Haq, A. Zeb, Z. Lei, and D. Zhang, "Forecasting daily stock trend using multi-filter feature selection and deep learning," *Expert Syst. Appl.*, vol. 168, no. December 2020, p. 114444, 2021, doi: 10.1016/j.eswa.2020.114444.

[68] A. Kanwal, M. F. Lau, S. P. H. Ng, K. Y. Sim, and S. Chandrasekaran, "BiCuDNNLSTM-1dCNN — A hybrid deep learning-based predictive model for stock price prediction," *Expert Syst. Appl.*, vol. 202, no. August 2021, p. 117123, 2022, doi: 10.1016/j.eswa.2022.117123.

[69] M. C. Lee, J. W. Chang, J. C. Hung, and B. L. Chen, "Exploring the effectiveness of deep neural networks with technical analysis applied to stock market prediction," *Comput. Sci. Inf. Syst.*, vol. 18, no. 2, pp. 401–418, 2021, doi: 10.2298/CSIS200301002L.

[70] G. S. Atsalakis and K. P. Valavanis, "Surveying stock market forecasting techniques - Part II: Soft computing methods," *Expert Syst. Appl.*, vol. 36, no. 3 PART 2, pp. 5932–5941, 2009, doi: 10.1016/j.eswa.2008.07.006.

[71] G. S. Atsalakis, "S URVEYING S TOCK M ARKET F ORECASTING T ECHNIQUES - P ART I :," no. January 2013, 2014.

[72] H. Maqsood *et al.*, "A local and global event sentiment based efficient stock exchange forecasting using deep learning," *Int. J. Inf. Manage.*, vol. 50, no. December 2018, pp. 432–451, 2020, doi: 10.1016/j.ijinfomgt.2019.07.011.

[73] F. Jia and B. Yang, "Forecasting Volatility of Stock Index: Deep Learning Model with Likelihood-Based Loss Function," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/5511802.

[74] T. Yin, C. Liu, F. Ding, Z. Feng, B. Yuan, and N. Zhang, "Graph-based stock correlation and prediction for high-frequency trading systems," *Pattern Recognit.*, vol. 122, p. 108209, 2022, doi: 10.1016/j.patcog.2021.108209.

[75] M. Y. Day and C. C. Lee, "Deep learning for financial sentiment analysis on finance news providers," *Proc. 2016 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2016*, no. 1, pp. 1127–1134, 2016, doi: 10.1109/ASONAM.2016.7752381.

[76] W. Long, Z. Lu, and L. Cui, "Deep learning-based feature engineering for stock price movement prediction," *Knowledge-Based Syst.*, vol. 164, pp. 163–173, 2019, doi: 10.1016/j.knosys.2018.10.034.

[77] E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," *Expert Syst. Appl.*, vol. 83, pp. 187–205, 2017, doi: 10.1016/j.eswa.2017.04.030.

[78] C. Wang, H. Liang, B. Wang, X. Cui, and Y. Xu, "MG-Conv: A spatiotemporal multi-graph convolutional neural network for stock market index trend prediction," *Comput. Electr. Eng.*, vol. 103, no. May, p. 108285, 2022, doi: 10.1016/j.compeleceng.2022.108285.

[79] G. Bathla, R. Rani, and H. Aggarwal, "Stocks of year 2020: prediction of high variations in stock prices using LSTM," *Multimed. Tools Appl.*, 2022, doi: 10.1007/s11042-022-12390-5.

[80] Ishwarappa and J. Anuradha, "Big data based stock trend prediction using deep CNN with reinforcement-LSTM model," *Int. J. Syst. Assur. Eng. Manag.*, 2021, doi: 10.1007/s13198-021-01074-2.

[81] J. Nayak, P. B. Dash, B. Naik, S. Mohapatra, and A. R. Routray, "Deep Learning-Based Trend Analysis on Indian Stock Market in COVID-19 Pandemic Scenario and Forecasting Future Financial Drift," *J. Inst. Eng. Ser. B*, 2022, doi: 10.1007/s40031-022-00762-2.

[82] H. Goel and N. P. Singh, "Dynamic prediction of Indian stock market: an artificial neural network approach," *Int. J. Ethics Syst.*, vol. 38, no. 1, pp. 35–46, 2022, doi: 10.1108/IJOES-11-2020-0184.

[83] A. Boru İpek, "Stock price prediction using improved extreme learning machine methods during the Covid-19 pandemic and selection of appropriate prediction method," *Kybernetes*, 2022, doi: 10.1108/K-12-2021-1252.

[84] S. Sinha, S. Mishra, V. Mishra, and T. Ahmed, "Sector influence aware stock trend prediction using 3D convolutional neural network," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1511–1522, 2022, doi: 10.1016/j.jksuci.2022.02.008.

[85] H. S. Sim, H. I. Kim, and J. J. Ahn, "Is Deep Learning for Image Recognition Applicable to Stock Market Prediction?," *Complexity*, vol. 2019, 2019, doi: 10.1155/2019/4324878.

[86] D. Tashiro, H. Matsushima, K. Izumi, and H. Sakaji, "Encoding of high-frequency order information and prediction of short-term stock price by deep learning," *Quant. Financ.*, vol. 19, no. 9, pp. 1499–1506, 2019, doi: 10.1080/14697688.2019.1622314.

[87] K. Liu, J. Zhou, and D. Dong, "Improving stock price prediction using the long short-term memory model combined with online social networks," *J. Behav. Exp. Financ.*, vol. 30, p. 100507, 2021, doi: 10.1016/j.jbef.2021.100507.

[88] F. Liu, P. Qin, J. You, and Y. Fu, "Sparrow Search Algorithm-Optimized Long Short-Term Memory Model for Stock Trend Prediction," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/3680419.

[89] G. Ma, P. Chen, Z. Liu, and J. Liu, "The Prediction of Enterprise Stock Change Trend by Deep Neural Network Model," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/9193055.

[90] X. Teng, X. Zhang, and Z. Luo, "Multi-scale local cues and hierarchical attention-based LSTM for stock price trend prediction," *Neurocomputing*, vol. 505, pp. 92–100, 2022, doi: 10.1016/j.neucom.2022.07.016.

[91] J. J. Han and H. jung Kim, "Prediction of Investor-Specific Trading Trends in South Korean Stock Markets Using a BiLSTM Prediction Model Based on Sentiment Analysis of Financial News Articles," *J. Behav. Financ.*, vol. 0, no. 0, pp. 1–13, 2021, doi: 10.1080/15427560.2021.1995735.

[92] H. Zheng, H. Wang, and J. Chen, "Evolutionary Framework with Bidirectional Long Short-Term Memory Network for Stock Price Prediction," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/8850600.

[93] U. Gupta, V. Bhattacharjee, and P. S. Bishnu, "StockNet—GRU based stock index prediction," *Expert Syst. Appl.*, vol. 207, no. March 2021, p. 117986, 2022, doi: 10.1016/j.eswa.2022.117986.

[94] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," *Eng. Appl. Artif. Intell.*, vol. 115, 2022, doi: 10.1016/j.engappai.2022.105151.

[95] D. Muller, I. Soto-Rey, and F. Kramer, "An Analysis on Ensemble Learning optimized Medical Image Classification with Deep Convolutional Neural Networks," *IEEE Access*, no. May, pp. 1–1, 2022, doi: 10.1109/access.2022.3182399.

[96] A. B. Zhao and T. Cheng, "Stock return prediction: Stacking a variety of models," *J. Empir. Financ.*, vol. 67, no. April, pp. 288–317, 2022, doi: 10.1016/j.jempfin.2022.04.001.

[97] D. Xiao and J. Su, "Research on Stock Price Time Series Prediction Based on Deep Learning and Autoregressive Integrated Moving Average," *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/4758698.

[98] F. Ronaghi, M. Salimibeni, F. Naderkhani, and A. Mohammadi, "COVID19-HPSMP: COVID-19 adopted Hybrid and Parallel deep information fusion framework for stock price movement prediction," *Expert Syst. Appl.*, vol. 187, no. September 2021, p. 115879, 2022, doi: 10.1016/j.eswa.2021.115879.

[99] K. Yadav, M. Yadav, and S. Saini, "Stock values predictions using deep learning based hybrid models," *CAAI Trans. Intell. Technol.*, vol. 7, no. 1, pp. 107–116, 2022, doi: 10.1049/cit2.12052.

[100] H. Wang, J. Wang, L. Cao, Y. Li, Q. Sun, and J. Wang, "A Stock Closing Price Prediction Model Based on CNN-BiSLSTM," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/5360828.

[101] S. R. Polamuri, D. K. Srinivas, and D. A. Krishna Mohan, "Multi-Model Generative Adversarial Network Hybrid Prediction Algorithm (MMGAN-HPA) for stock market prices prediction," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7433–7444, 2021, doi: 10.1016/j.jksuci.2021.07.001.

APPENDIX

Table 7: Comparison of Stock Prediction Algorithms using Deep Learning

Study	Dataset	Input Feature	Methods	Performance Measurement
Sharaf *et al.* [33]	Nasdaq Stock Exchange	OHLC price, Volume	LSTM, CNN, SVM, Linear Regression, Logistic Regression, K-Neighbors, Decision Tree, Random Forest, Stacked-LSTM, Bidirectional-LSTM	RMSE=0.15 MSE=0.22 MAE=0.089 MAPE=39.9%
Lee and Yoo [35]	Korea Exchange, New York Stock Exchange, Nasdaq Stock Exchange	OHLC price, Volume	DNN	Hit Ratio=0.49
Tian *et al* [36]	New York Stock Exchange	OHLC price, Volume	LSTM and LightGBM	RMSE= 596.04 Accuracy = 63.9%
Li *et al.* [37]	New York Stock Exchange, Nasdaq Stock Exchange	OHLC price, Volume	LSTM, RNN, GRU	MSE=0.0059 RMSE=0.0745
Thakkar and Chaudhari [22]	Korea Exchange, Shanghai Stock Exchange	OHLC price, Volume, technical indicatorL	LSTM	MAPE= 0.0706 MSE= 33.9963
Rather [39]	National Stock Exchange	Close price	LSTM	RMSE= 0.375
Carta *et al.* [40]	New York Stock Exchange, Frankfurt Stock Exchange	Close price	An ensemble of Deep Q-learning	
Rezaei, Faaljou and Mansourfar [41]	New York Stock Exchange, Frankfurt Stock Exchange, Tokyo Stock Exchange	Close price	CEEMD-CNN-LSTM EMD-CNN-LSTM	RMSE=13.76 MAE=10.58 MAPE=0.536
Chen, Wu and Wu [42]	Shanghai Stock Exchange	Close price	K-Means LSTM	MSE= 0.0022 MAE= 0.0328
Liu and Ma [43]	Nasdaq Stock Exchange Bombay Stock Exchange Hong Kong Stock Exchange Shanghai Stock Exchange New York Stock Exchange Taiwan Stock Exchange	Close price	Quantum ANN	*Error* = 0.10163
Tao *et al.* [44]	Shenzhen Stock Exchange	Close price	ConvLSTM	MSE= 1.0547 MAE= 0.7060 MAPE=0.0123
Q. Liu *et al.* [45]	Shanghai Stock Exchange Shenzhen Stock Exchange	Close price, Fundamental Data	Deep Learning Neural Network	Accuracy=55.46%
Zhang and Lou [46]	Shanghai Stock Exchange	OHLC price, Volume, Fundamental Data	Backpropagation NN	Accuracy= 73.29%
Chen *et al.* [28]	Shanghai Stock Exchange, Hong Kong Stock Exchange	Close Price, Volume, News, technical indicator	Bi-LSTM	DA=0.608 MCC= 0.1072
Wu *et al.* [30]	Shanghai Stock Exchange	OHLC Price, Volume, Sentiment, technical indicator	S_I_LSTM	MAE= 2.386835 MSE= 7.271708 RMSE=2.6966
Shilpa and Shambhavi [31]	National Stock Exchange	Stock price data, Sentiment, technical indicator	NN, DBN, SIWOA	MAE=0.21
Li *et al.* [47]	Shanghai Stock Exchange	OHLC price, Volume, Sentiment, Fundamental data	LSTM	DA=0.624 MCC=0.4472

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Song, Lee and Lee [57]	Korea Exchange	Price and technical indicator	DNN	Accuracy=84.80%
X. Liu *et al.* [24]	Shanghai Stock Exchange Shenzhen Stock Exchange New York Stock Exchange	OHLC price, Volume dan technical indicator	BP Neural Network ISSA	MAE=3.0760 MAPE%=0.0934 RMSE=3.1980
Wu, Wang and Wu [25]	Shanghai Stock Exchange Shenzhen Stock Exchange	OHLC price, amount, and technical indicator	EML	Accuracy=77.82%
Chen, Jiang, *et al.* [26]	Shanghai Stock Exchange	OHLC price and technical indicator	GC-CNN	Accuracy = 52.20%
C. Ma *et al.* [59]	Shanghai Stock Exchange Shenzhen Stock Exchange	Close price and Volume	LSTM	RMSE= 0.074 MSE=0.006 MAE= 0.059
Bathla, Rani and Aggarwal [79]	National Stock Exchange, Bombay Stock Exchange, Nasdaq Stock Exchange, New York Stock Exchange, Tokyo Stock Exchange	Close price	LSTM	MAPE for each stock exchange: 3.89; 1.21; 3.01; 1.19; 2.03; and 0.86
Nayak *et al.* [81]	National Stock Exchange	Close price	LSTM	MAPE=0.0030, MAE=0.0065, RMSE=0.0073
Chen *et al.* [28]	Shanghai Stock Exchange, Hong Kong Stock Exchange	Close price, Volume, technical indicator and sentiment	Bi-LSTM	DA=0.613 MCC=0.1165
Gupta, Bhattacharjee and Bishnu [93]	National Stock Exchange	Open price	GRU	RMSE=0.0896 MAE= 69.9396 MAPE= 0.8203
Li and Pan [61]	New York Stock Exchange	Close price, sentiment	Ensemble Deep Learning	MSE=186.32 Precision=60%
Zhao and Cheng [96]	New York Stock Exchange	Economic and macroeconomic indicators	Ensemble Learning	Accuracy=79.03%
Xiao and Su [97]	New York Stock Exchange	Close price	ARIMA LSTM	MSE=0.01 RMSE=0.319 MAE=0.248
Ronaghi *et al.* [98]	New York Stock Exchange	OHLC price and sentiment	CNN-BLSTM	Accuracy= 66.48%
Yadav, Yadav and Saini [99]	New York Stock Exchange	Close price	Fast RNN - CNN - BiLSTM	RMSE= 0.14932
Kumar *et al.* [34]	New York Stock Exchange, Nasdaq Stock Exchange	OHLC price and Volume	Generative Adversarial Network (GAN)	Directional Accuracy=64.58%
Diqi, Hiswati and Nur [38]	Indonesia Stock Exchange	OHLC price and Volume	Generative Adversarial Network (GAN)	MAE= 0.020665 R2 Score = 0.811