Journal of Statistics, Computing and Interdisciplinary Research
ISSN (P): 2707-7101

J. Stat. Comp. & Interdis. Rese.
Volume 7(1) (2025), 1-25
http://scir.wum.edu.pk/index.php/ojs

Predicting Stock Exchange Closing Prices using Machine Learning Techniques

Abdul Qadeer Mudasar[1], Fatima Bukhari[1], Ruqqia Talib[2] *, Imran Khurshid[1], Muhammad Ahsan Jamil[3], Khadija Naqvi[4]

[1]Department of Computing, National University of Modern Languages, Multan, 60000, Pakistan.

[2]Department of Information Technology, Bahauddin Zakariya University, Multan, 60000, Pakistan.

[3]Institute of Computing, MNS University of Agriculture, Multan, 60000, Pakistan.

[4]Department of Economics, The Islamia University of Bahawalpur, 63100, Pakistan.

***Corresponding Author:** Ruqqia Talib, ruqqiatalib@gmail.com

Abstract

Predictive analysis powered by machine learning is gaining momentum, especially when it comes to using historical data to anticipate future trends. One area where this is particularly useful—but also quite challenging—is in forecasting stock market closing prices. The complexity stems from the stock market's inherently unpredictable and non-linear nature. In this study, we present a tailored approach that combines two powerful machine learning models: Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM). By leveraging historical stock data, our goal is to predict closing prices with greater accuracy. Our dataset focuses on four major companies listed on the Pakistan Stock Exchange: Amreli Steel Limited (ASTL), Meezan Bank Limited (MEZL), United Bank Limited (UBL), and Oil and Gas Development Company (OGDC). For each, we've used financial indicators such as Open, High, Low, Average, and Last Day Close Price (LDCP) as input features for our models. The main objective of our research is to evaluate how well ANN and LSTM models perform in forecasting stock closing prices. We conduct a detailed analysis, applying rigorous testing methods to assess both accuracy and effectiveness. Our findings aim to contribute to smarter stock market predictions by offering practical insights. More reliable forecasts can empower investors and market stakeholders to make informed decisions and better manage risk. Importantly, the approach we've developed goes beyond traditional forecasting methods by integrating

advanced machine learning techniques. These models are capable of identifying complex patterns in financial data, potentially leading to more accurate predictions. Additionally, the use of multiple input variables allows for a well-rounded analysis of the many factors that influence stock prices.

Keywords: *Embedded systems; Java bytecode optimization; Storage optimization; Interpreter optimization; Superinstructions; Java instruction set.*

Introduction

Undoubtedly, the stock market holds a pivotal significance in a nation's overall economic status, with its influence transcending international boundaries. Accurately predicting stock market movements, including closing prices and trends, is of great significance for investors, financial institutions, and policymakers. With the global nature of investments and the interconnectedness of financial markets, precise predictions enable informed decision-making, strategic planning, and risk management. Consequently, the development and application of robust predictive models and algorithms in stock market analysis have gained immense importance in recent times.

The infrequent economic assets could be sufficiently apportioned to the most cost-effective project and actions [11]. Investors put their assets into the stock market with the goal of earning higher profits. Using an accurate algorithm to predict stock closing prices and the various economic and other factors affecting its unpredictability can help individuals benefit from the financial market. [11]. Due to the financial stock market’s turbulent and non-linear data form, forecasting stock exchange data can be difficult. The forecasting of stock market trends has been a challenging and worthwhile research issue. Data is challenging because noise, not stasis, exists. Information is crucial because information can help decision-makers reach key conclusions.[19] The application of machine learning and deep learning is extensive in data prediction and forecasting. Conventionally, different algorithms have been suggested and forecasting the stock price of the organization. [32] The first technical analysis of this process using the historical price of the stock like high, low, open, average, and last day closing is normally used. Secondly, the process is qualitative which properly on the external factors like market situation, company profile other factors like political, country conditions, blog, and social media. [33]

Machine learning falls within the realms of artificial intelligence and computer science, encompassing the acquisition of data learning from various perspectives. Machine learning

provides difficult patterns and analysis of the data to learn for different purposes like forecasting. Current business models are facing many problems dealing with large data which is difficult to understand for humans. Machine learning algorithms helps to understand and manipulate large data within a short time period which is impossible for human beings. Machine learning helps remove the bias from the data which problem normally faced by humans.

Machine learning algorithms contribute to enhancing data accuracy compared to human-made predictions, thereby yielding business advantages. Machine learning algorithm helps future decisions about business plans. To create a machine learning algorithm, some main steps are data collection, preprocessing of the collected data, and designing the model which performs more results on the data. Different machine learning algorithms are used on different data like Labeled data, unlabeled data, and data learning from the environment. Supervised machine learning algorithms primarily rely on labeled data, unsupervised machine learning algorithms utilize predominantly unlabeled data, and reinforcement machine learning often draws knowledge from learning environments. Some popular machine learning algorithms have been utilized in previous works, including [32][33][31][4][11]: Linear Regression, Back Propagation, Partial Least Square, Convolutional Neural Network, Sequential Minimal Optimization, Long Short-Term Memory, and Multi-layer Perceptron.

Past examinations have inspected different free factors as contributions to Artificial Intelligence models. The proposed framework in this study utilizes a portion of similar free factors as past work and presents new ones. The calculations used in the proposed framework include:

- Artificial Neural Network
- Long Short-Term Memory

The proposed algorithm is customized for forecasting the closing price, and its accuracy is compared to other algorithms to obtain optimal results. The corpus used in this study will consist of Stock Exchange data obtained from a freely available online resource called Kaggle [14]. This data includes information on the opening, high, low, average, close, and previously close prices of several listed companies, which will be analyzed in this research:

- Amreli Steel Limited (ASTL)
- Meezan Bank Limited (MEZL)
- United Bank Limited (UBL)

- Oil and Gas Development (OGDC)

In the proposed system following the company historical data is used for the forecasting of the data. Before applying the model to the corpus, the first step is preprocessing. In preprocessing data is analyzed and maintained to fit our customized model for forecasting. Then we create a new variable which is used as an independent variable. Normalization of the data will be applied for getting better-processed data. Then splitting will be done and applying the model to the processed data and different visualization is generated. The following question is answered in this research.

- How will Our proposed model work?
- How does the proposed model forecast the closing price?
- What will the accuracy and success of our proposed model?

Finally got our aims of the proposed system.

- To design a model
- Forecasting the closing price of companies listed on the Stock Exchange
- Comparing our proposed model's accuracy

Related Work

Researchers commonly utilize machine learning and deep learning techniques, with prominent algorithms, as a foundation for predicting stock market data. Furthermore, this section also describes the current model used in the proposed work. Machines and deep learning are quite helping another area of study in every case of life. Machine learning is categorized into three primary types: Reinforcement Learning, Supervised Learning, and Unsupervised Learning. The procedure of machine learning contains the following step to follow problem understanding, data preprocessing modeling of data, and fitting data to the model. Normally these steps are following those described in the last lines.

Accuracy in the prediction of the stock market is difficult to perform due to the behavior of a financial stock. According to [32], the author's description of intelligence prediction of the stock market data is efficient and more help full. The author devised an approach involving Artificial Neural Network and Random Forest models applied to data from five distinct companies. ANN can find hidden features through a self-learning process. Random Forest is applicable to both regression and classification. [32] In some previous papers [31], they used linear regression, time series forecasting, exponential smooth, and month moving average. They used the dataset of Amazon, GOOGLE, AAPL, and yahoo finance stock market trend

for upcoming months. The financial market is increasing continuously because people are more interested in stock market investment. Artificial Intelligence helps predict financial crises and investments in the stock market. [31][16]

Stock market predictions are fascinating and demanding research in the stock market. According to the author's [16] much of the previous research has taken on the different stock market datasets and different famous and non-famous machine learning algorithms. Author's works on predicting the future trend of the markets: Karachi, London, New York. The stock market is the place where different companies and people invest their large investments. Future trends help investors to invest in the correct trend and correct time. Principle component analysis (PCA) is used by [16], the authors for predicting future price trends. KNN provides better accuracy than in these datasets on the dependence of the conditions of data. [16]
Review work is done by [29] which describes the stock market trends and different machine learning algorithms used by the different authors in his work. According to him, the stock market contains the relay in data. But machine learning is more efficient than other techniques. A systematic literature process is used to find out the related paper to work. ANN, SVM, GA, and many combined models of these algorithms to generate the hybrid model and achieve the goals. Find out some unique limitations in the works, some future works, and some new areas of work in the stock exchange. According to the author's study of alternative genetic algorithms, these other two ANN and SVM can be used for prediction. The Conclusion of the author's study [29] is that financial investment needs machine learning.

Some of the authors work on the stock market in different styles. Like the impact of the day of the week on the market, there are also weekend effects and specific occasions that can influence the stock market. Additionally, the monthly effect is considered when analyzing the data. Pakistan Stock Exchange Data (PSX) some dependent on the weekly and monthly effects on the market like some religious days and some holidays. The Box-Jenkins (ARIMA) method is considered the most suitable approach for examining the time series model against historical data. Additionally, multiple regression analysis is applied, incorporating models such as ARCH, least squares, and EGARCH-in-mean. [4] The research result shows that Friday is a more effective market for the whole week in the Karachi stock exchange. December and March month is a hot month for different companies on the Karachi stock exchange. [4][26] Previous studies have highlighted the existence of herding behavior in the dataset of the Pakistan Stock Exchange. [26] The behavior of the investor is mainly dependent on the stock

market investment. The factors are that some investors do not want to invest in small stocks and some in large stocks. Some of them do not want to invest in the new companies and some will work with only the top brands only. A study on this is done by [26] in which they used different previous techniques about the stock market Harding behavior and conclude that this will help with estimation.

Prior studies have illuminated the existence of herding behavior within the dataset of the Pakistan Stock Exchange. [15] The authors used the fiver year dataset of the pharmaceutical companies which impacted the payout. For the finding of this work, they used backward linear regression and correlations analysis on the data and the association between the variables to predict the payout. According to [15] 31.90% of payout depends on the variable: audit type, growth opportunities, and liquidity of the pharmaceutical dataset. The remaining variable like risk and firm size already listed above mentioned is the effect on the payout condition according to the decision. [2] Industrial forecasting is helpful according to finding the yearly expenditure of any organization. Forecasting is a thing that talks about something of the future that may be possible in the future. So that way of predicting or forecasting the industrial financial data is effective for this work. Every year the organization expenses billions of monies on the expenses.

Operational resources in a financial workplace are also defined as expenditure. The expenses have been made for the linear and predictive growth of the finances. The researchers utilize Long Short-Term Memory (LSTM), a type of recurrent neural network (RNN) in the field of deep learning. This is combined with the integration of specific statistical capabilities from IBM SPSS to facilitate forecasting. The dataset utilized is Pakistan's DGP, reflecting expenditure at prevailing prices. This model is anticipated to yield predictive outcomes with heightened accuracy compared to conventional statistical methodologies. [2][22]

Forecasting the big data of stock market prices become more famous using deep learning models. They [7] gather data on the Chinese stock market for two years. Model based on deep learning and feature engineering suggested by the author. The proposed research introduces a deep learning-driven model, which is integrated with the preprocessing of the stock market dataset and incorporates various feature engineering strategies to predict fluctuations in stock prices. They conducted a thorough analysis of the most popular machine learning models before coming up with their own model. [27]

The endeavor to utilize machine learning for predicting stock index prices has been a central focus of research efforts. The intricate characteristics of stock market data, characterized by its non-linear and intricate behavior, pose a substantial challenge. In response, deep learning models have arisen as a compelling solution for handling such intricate data structures. Scholars have unveiled an amalgamated deep learning model that combines several tiers of convolutional neural networks and bidirectional long short-term memory units. This integrated model has shown a substantial 9% enhancement in predictive accuracy when juxtaposed with single-layer deep learning models. Furthermore, it significantly surpasses a support vector machine regressor model, as confirmed through evaluation on the S&P 500 grand challenge dataset. In order to tackle complexities and reduce overfitting, researchers have introduced a range of modifications to single and multiple-layer deep learning models. These adaptations incorporate diverse convolutional neural network kernel sizes and bidirectional LSTM units. The mounting effectiveness of deep learning models within the financial domain has garnered substantial attention from both investors and researchers alike. [5][9]

Investors in the stock market with successful profiles can earn a maximum range of profit from the stock exchange with the selection of the trading suitable time and company. Two main fundamental statistical techniques for deciding are used. For analysis, two statistical approaches, namely fundamental analysis and technical analysis, are utilized. Last few years machine learning to enhance in field of predicting investors to enhance investment. Machine learning models are used for stock predictions. Within the realm of machine learning models, there are neural networks, support vector machines, and genetic algorithms. Multiple multiclass classification methods utilizing neural networks were used by the authors [6]. They evaluated and contrasted the One-Against-One (OAO) and One-Against-All (OAA) approaches with the conventional neural network. A data selection strategy is utilized, along with many alternative data preparation techniques. The final experimental findings demonstrate that the OAA technique's multi-binary classification beats other methods. [6]
Table 1 describes a comparison of other works that are done already in this field and some sort of description is also mentioned above:

Year	**Models**	**Best Model**	**Accuracy**	**Dataset**	**Limitation**	

2023	Multiclass classification model	XGBoost and AdaBoost	Accuracy 65%	Tehran change Stock Exchange	Different technique joined to get more accurate result	[30]
2023	The assessment of machine learning methodologies for predicting stock market trends has attracted considerable attention					[8]
2023	ARIMA,SVR,LSTM, and XG-Boost	LSTM		HDFC Pharma and Sun		[28]
2023	An Examination of Regularization in Machine Learning Models					[20]
2022	LSTM, GRU, CNN	LSTM	RMSE : 9.4338 MAPE : 0.5826	Nepal Stock Exchange	Hybrid model can be created using proposed model	[23]
2022	GORU and variational auto-encoder (VAE)	Combination of both	Precision : 0.57289 Recall : 0.97063	Stocknet-dataset, Astock-dataset	Furthermore Lower memory and time complexity	[18]
2022	LSTM, Hybrid LSTM, ARIMA, CNN, Hybrid CNN				Review of models on stock pre-diction	[25]
2021	LSTM	LSTM	Accuracy : 63.59%,56.25%,and 57.95%	HDFC, Yes Bank, and SBI	Optimization technique can be used for further improvement	[1]
2021	Least-Squares Linear Regression Model	Least Squares Linear Regression Model	MAPE 1.367% and RMSE : 0.512	Bank of Am erica stock dataset	Day range prediction can be added for further accurate results	[10]
2021	CNN, RNN, LSTM, MLP, SVM	CNN		National Stock Exchange (NSE) of India	Addition of NLP and financial time series effect good on performance	[3]
2021	Survey of Stock Price and Forex					[13]

2021	LR, SVM, FFN, and RNN, Auditory Algorithm (AA)	Auditory Algorithm (AA)	p-value: 15.87	Oil and Gas Industry within the Nigerian Stock Exchange	Cryptocurrency prediction can be get using this approach	[21]
2020	Comprehensive examination of machine learning and statistical methodologies through a systematic review. techniques on stock data					[17]
2020	Artificial Neural Network(ANN), Random Forest	ANN	RMSE: 1.10, 3.30, 1.28, 1.54, 0.42 MAPE: 1.07%, 1.09%, 0.89%, 0.70%, 0.77% MBE: 0.052, 0.076, 0.031, 0.013, 0.015	Nike, Johnson and Johnson, JP Morgan and Co. Goldman Sachs, ,Pfizer Inc.	More Variable like traded Volume can be used	[32]
2019	Sequential Minimal Optimization, Multi-Layer Perceptron, and Partial Least Square	MLP	83.37%	Thai Stock Exchange (SET)	Anticipating Long-Term Stock Price Trends	[33]
2019	Utilizing Excel for Time Series Forecasting, Three Month Moving Average, Linear Regression , and Exponential Smoothing	Exponential Smoothing better result	Standard Error 4.903824	Google, Amazon, AAPL Stock	Other Machine Learning algorithm can be used	[31]
2016	KNN, Support Vector Machines, and Naïve Bayes	KNN	45% 68% 57%	LSE, NYSE, KSE	Addition of Social Media Data in historical Data	[16]
2014					stock market feature can be compared with economic growth	[19]
2017	Systematic Literature Review			Different Organization Dataset	Hybrid Machine Learning Model	[29]
2020	MLP, CNN, LSTM	LSTM	RMSE: 35.49, 228.87, 49.5815 R: 0.98,0.97,0.98 MAPE: 0.009, 0.0079, 0.010	SP500 index, CSI300 index, Nikkei225 index	Combining the linear and nonlinear model for the new model in stock market prediction	[11]

2020	Multiple Linear Regress using dummy variables	ARIMA Model	R2 = 0.7672	Karachi Stock Exchange Data	Settlement period hypothesis and retail investor hypothesis can examine also	[4]
2017	The herding behavior model proposed by Christie and Huang		There is considerable dispersion between the stock returns of individual firms and the market index.	Normally market index does not herd but experience a negative return of 5%	Herding behavior of individual investor will be tested	[26]
2017	Correlation Analysis, Backward Multiple Linear Regression (BMLR)	BMLR	32%	dependent on the audit type, liquidity, growth opportunities	All over world pharma companies' data can be used	[15]
2019	Artificial Neural Networks, Neural Networks, Long Short-Term Memory,	LSTM(SELU)	91%	Pakistan GDP	Furthermore, other activation function can be used	[2]
2016	Sentiment Analysis, Random Forest, Logistic Regression	Random Forest	70.18%	Twitter Sen timent Data in N-gram and Word2vec Form	Furthermore, Other social media data can be used	[22]
2017	SVM, KNN	Hybrid Framework	MAPE=0.21 to 0.31 In different prospectus ranges between them	Chinese Companies Historical Dataset	Correlation weighted method to be considered	[7]
1997	Long Short-Term Memory	Long Short-Term Perform Very well given situations		Theoretically analyzed	Can be performed on any dataset practically	[12]
2018	RAND, ARIMA, Random Forest, TSLDA, HAN	HAN	58.23%	Chaotic data	Furthermore, Independent Variable can be created for better results	[34]
2020	LR, SVM, MLP, LSTM	LSTM	93.25%	Chinese Stock Market data (2 year)	Social Media and news information can be added	[27]

2019	CNN, Bi-Directional LSTM	LSTM	Mean Test Score= 0.00034 Mean Train Score= 0.00027	S&P 500 grand challenge dataset	Combining this model to the sentiment and news data for better prediction	[9]
2017	Wavelet Transformation, Stacked Autoencoders, LSTM	LSTM	LSTM Perform better results in different level of testing	CSI 300 Index, Nifty 50 Index, Hang Seng Index, Nikkei 225 Index	Acquiring profound education takes time through different methodologies, and there's a necessity to emphasize avenues involving GPU-based accelerated and contrasting computing methods for mastering intricate learning	[5]
2019	Levenberg-Marquardt, Scaled Conjugate, Bayesian Regularization	Levenberg-Marquardt	MAPE = 99.981 MSE = 99.983	Tick data	Recurrent Neural Network used for further improvement	[24]
2016	Neural Network, One-Against One Neural Network, One-Against-All-Neural Network	One Against All-Neural Network	Average = 82%	10 different datasets	Fundamental and technical data can be applied further	[6]

Previous details we analyzed the literature review to figure out the critical points and missing research areas that can be reviewed further. Some gaps are clarified that can be used for further research. The novelty of the proposed work is to use of a machine learning algorithm with new created independent variable from which we enhance the prediction of previous practical work and worked on Pakistan stock exchange data which is different from data from other work.

Proposed Methodology

Most of the research was previously conducted in the stock exchange to predict the concluding price for the following day using numerical analysis, textual analysis, and practical analysis. Stock closing price prediction gets fame due to the addition of the machine learning field for predicting the price, high-value days, and many other predicting tasks in this area. Many authors use machine learning techniques for predicting stock predictions.

Machine learning helps very well in the latest research to provide a better solution and prediction of data in less time. Due to machine learning, we can make predictions using historical data. We can also say Machine Learning (ML) makes us better at using predictions. We are describing the proposed architecture. Previously we described the literature about our work then we make a comparison table about previous work done and draw an existing architecture. The next proposed architecture contains the Description of data, New Variable, Algorithm, Flow Chart, proposed High-Level Architecture of techniques, and description of the proposed methodology. In this research, we conduct a practical analysis of the stock exchange data of the different companies. We selected four companies' Stock Exchange data which are Amreli Steel Limited (ASTL), Meezan Bank Limited (MEZL), United Bank Limited (UBL), and Oil & Gas Development (OGDC). The historical data for the above-listed companies has been collected from Kaggle [14]. The dataset contains almost 14 years of data from 12-1-2002 to 12-5-2018 of ASTL, MEZL, UBL, and OGDC. The stock contains the following columns Date, Open, Low, High, Average, Last Day Closing Price, and Close. We used Open, High, Low, and Close averages used for the prediction of the value. Below Table 2 describes the dataset statistics and data used for training and testing.

Table 2. Description of Dataset

Company	Dataset	Training Dataset	Testing Dataset
ASTL	12/01/2015 - 5/4/2018	80%	20%
MEZL	3/28/2002 - 5/4/2018	80%	20%
UBL	7/25/2005 - 5/4/2018	80%	20%
OGDC	1/19/2004 - 5/4/2018	80%	20%

New Variable

A new average of Open High Low Close (OHLC) average has been created for the forecasting of the closing price of the stock exchange. Description of a new variable shown in Figure 1.

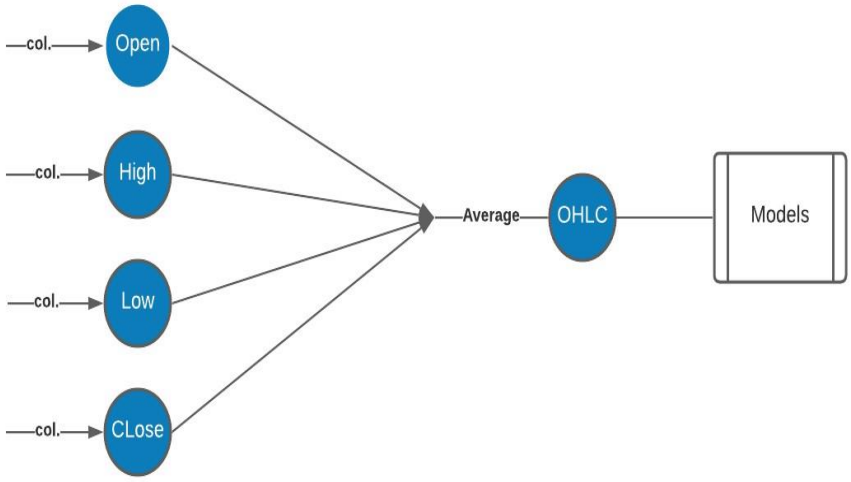

Figure 1 New Variable

The proposed methodology contains the artificial neural network and long short-term memory. To understand the artificial neural network first we have to understand the neural network. A neural network designed on the human brain contains many artificial neurons, which can be used on arranged in a sequence of layers. The layers contain an input layer, a hidden layer, and the output layer. Below equation (1) represents the artificial neural network. The equation (1) defines the input layer values putting that called weight and bias.

$$Y = W1 * X1 + ... + Wn * Xn + b \qquad (1)$$

where:

Y = Output/Dependent Variable

W1..Wn = The weights

X1..Xn = The input/independent variable

The first part of an LSTM network is the input gate, which controls the flow of information into the network. An LSTM (Long Short-Term Memory) cell is specially designed to handle sequences of data and remember important information over time. It does this through a combination of gates that control the flow of information.

The first is the input gate. This gate looks at both the current input and the previous output and decides how much of the new information should be allowed into memory. In other words, it filters what's important right now and passes that along to the next step.

Next is the forget gate. This part is crucial because it determines what information from the past should be kept or discarded. It also considers the current input and the previous output to decide which parts of the cell's memory are no longer relevant and can be forgotten. Then

comes the output gate, which controls what the LSTM cell sends out as output. Like the other gates, it looks at the current input and previous output, then decides how much of the updated memory should be shown as the output for the current step. At the core of the LSTM is a memory cell, which stores information over time. This memory is updated by the input and forget gates, allowing the LSTM to hold onto valuable information from earlier in the sequence while discarding what's no longer needed. Thanks to this smart gating system, LSTMs are excellent at managing long-term dependencies in data. They can remember context from earlier in a sequence and use it to make better predictions or decisions, something that makes them incredibly effective for tasks like language modeling, time-series forecasting, and more.

$$k_t = \sigma(W_j[h_t{-}_1, c_t] + b_j) \qquad (2)$$

In this context, b_i refers to the bias term for the input gate, c_t represents the input at the current time step, h_{t-1} signifies the hidden state from the preceding time step, W_j stands for the weight matrix for input data, and σ indicates the sigmoid function.

$$F_l = \sigma(W_d\,[h_l{-}_1, x_l] + b_d\,) \qquad (3)$$

In this context, b_d denotes the bias term for the forget gate, x_l signifies the input at the current time step, h_{t-1} represents the hidden state from the preceding time step, W_d stands for the weight matrix for the forget gate, and σ indicates the sigmoid function.

$$C_l = tanh(W_g[h_l{-}_1, x_l] + b_g) \qquad (4)$$

In the given context, b_g represents bias term for the candidate cell state, x_l stands for input at the present time step, h_{l-1} signifies hidden state from the preceding time step, W_g denotes weight matrix for the candidate cell state, and tanh is indicative hyperbolic tangent function.

The equation (5) is the combination of the equation (4),(3), and (2).

$$Cm = fm * Cm{-}1 + im * Cm \qquad (5)$$

$$O_l = \sigma(W_m[h_l{-}1, x_l] + b_m) \qquad (6)$$

In this context, b_m represents the bias term for the output gate, x_l denotes the input at the current step, h_{l-1} signifies the hidden state from the previous time step, W_m stands for the weight matrix for the output gate, and σ indicates the sigmoid function.

And the last equation (7) is the output layer equation for the LSTM.

$$H_m = o_m * tanh(C_m) \qquad (7)$$

Algorithm 1 Proposed Procedure for Artificial Neural Network

Input: OHLC Average of Dataset

Output: Predicting Closing Price of Stock Exchange

***Initialization*:**

Importing Dataset

Prepossessing the Dataset

if (Dataset == Null) **then**

Remove Values

end if

Feature Extraction

Splitting the Dataset

Reshaping in 2D Matrix

Initializing the Artificial Neural Network Layers

Adding Dense Layer

Compiling Model

Fitting Model with 150 epochs having 4 batch-size

Testing the Model using Standard Error Techniques Predicted Results

---

Algorithm 2 Proposed Procedure for Long-Short Term Memory

Input: OHLC Average of Dataset

Output: Predicting Closing Price of Stock Exchange

Initialization:

Importing Dataset

Prepossessing the Dataset

if (Dataset == Null) **then**

Remove Values

end if

Feature Extraction

Splitting the Dataset

Reshaping in 3D Matrix

Initializing the Long-Short-Term Memory

Adding Layers

Compiling Model

Fitting Model with 100 epochs having 32 batch-size

Testing the Model using Standard Error Techniques

Predicted Results

The proposed algorithm 1 entails training an Artificial Neural Network (ANN) to predict stock exchange closing prices using OHLC average data. The process begins with data preprocessing, which includes handling missing values and extracting meaningful features from the dataset. Once the data is cleaned and prepared, it is split into training and testing sets. This division sets the stage for building and training an Artificial Neural Network (ANN). The ANN used in this study consists of a dense layer and is trained using 150 epochs with a batch size of 4. After training, the model's performance is evaluated using metrics such as standard error to determine how well it predicts closing prices on unseen data. Once validated, the trained ANN can be used to forecast future stock exchange closing prices based on new inputs. In parallel, Algorithm 2 details the process for using a Long Short-Term Memory (LSTM) network to perform the same forecasting task. Similar to the ANN approach, this begins with importing and preprocessing the data, which includes managing missing values and extracting the OHLC (Open, High, Low, Close) average. After preprocessing, the dataset is divided into training and testing sets and reshaped into a 3D matrix—necessary for the LSTM model to process sequential data effectively. The LSTM model is then initialized, configured with multiple layers, and compiled. It is trained using 100 epochs and a batch size of 32. As with the ANN, the model's predictive performance is assessed using standard errors on the test data. A well-trained LSTM model can then be applied to predict future closing prices using new OHLC input values.

The complete architecture of our approach is visually represented in Figure 2. It outlines each step, beginning with the import of a CSV dataset containing stock data for listed companies. Preprocessing involves cleaning the data, handling any missing values, and calculating the OHLC average—an essential feature for both models. After this, data normalization is applied using the Standard Scaler method from the SciPy library in Python. The dataset is then split into training and testing sets. For the LSTM model, the final preprocessing step involves reshaping the input from a 2D format into a 3D structure to accommodate the LSTM layer's requirements. Unlike LSTM, the ANN model does not require reshaping. The outcomes of both approaches are further analyzed and discussed in the Results section.

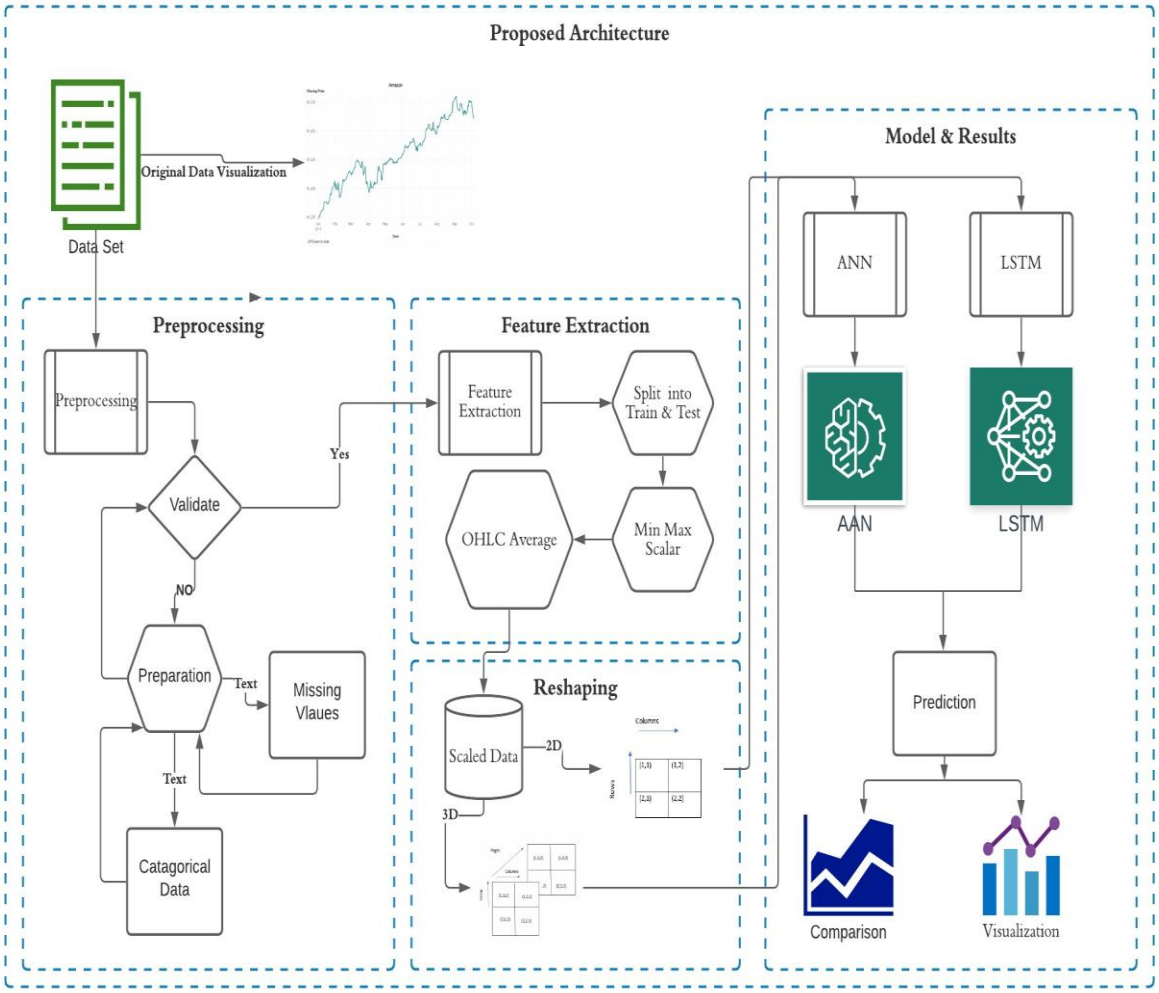

Figure 2 *Detailed Technical Description of Proposed Work*

Analysis & Results

To evaluate the accuracy of our model's stock price predictions, we use several key performance metrics: Mean Absolute Percentage Error (MAPE), Mean Bias Error (MBE), and Root Mean Square Error (RMSE). These metrics allow us to quantify the difference between the predicted and actual values, helping us understand how well the model is performing. MAPE provides a percentage-based error measure, making it easier to interpret. MBE indicates whether the model tends to overestimate or underestimate, while RMSE highlights the magnitude of prediction errors. Together, these metrics guide us in identifying and reducing errors, ultimately improving the reliability of our stock price forecasts.

RMSE [32] can be found using equation 8:

$$\text{RMSE}=\sqrt{\frac{\sum_{j=1}^{M}(T_j-P_j)^2}{M}} \qquad (8)$$

Here:

RMSE = Root Mean Square Error

J = variable j

M = total number of non-missing values

T_j = actual OHLC average price

P_j = predicted price

MBE [32] can be found using equation 9:

$$MBE = \frac{1}{M}\sum_{j=1}^{M}(T_j - P_j) \qquad (9)$$

Here:

MBE = Mean Bias Error

j = variable j

M= total size

T_j= actual OHLC average price

P_j= predicted price

MAPE [32] can be found using equation 10:

$$MAPE = \frac{1}{M}\sum_{j=1}^{M}\frac{(T_j - P_j)}{T_j} * 100 \qquad (10)$$

Here:

MAPE = Mean Absolute Percentage Error

j = variable j

M = total size

T_j = Actual OHLC average price

P_j = Predicted Price

The outcome of the proposed model is described in **Table 3**. We are constructing an evaluation table using the various error-finding approaches stated above, as well as the whole formula of the error rate-finding techniques of the models. Table 3 shows the differential value based on the data from the companies. Long Short-Term Memory (LSTM) clearly outperforms Artificial Neural Network (ANN).

Table 3. Result Evaluation table of the ANN and LSTM using MAPE, RMS and MBE

Company	**LSTM**			**ANN**		
	MAPE	**RMSE**	**MBE**	**MAPE**	**RMSE**	**MBE**
ASTL	2.8128	2.1168	-0.7241	3.9347	2.713684	1.5077
MEZL	3.1275	6.0227	-3.931204	5.0230	5.767712	3.56739
UBL	2.2192	4.7059	-1.053693	2.6264	5.217132	1.05155
OGDC	3.0984	4.2367	1.00602	5.3564	7.15731	-0.6996

The data presented in Table 3 reflects the performance of the LSTM and ANN models across different organizations. Reduced MAPE and RMSE values suggest heightened accuracy, while MBE values approximating zero indicate unbiased predictions. The evaluation encompasses three key performance metrics: MAPE, RMSE, and MBE, offering a comprehensive assessment of both LSTM and ANN models. Notably, in terms of accuracy and prediction precision, the LSTM model demonstrates superiority over the ANN model usually. When contrasted with the ANN model, the LSTM model exhibits lower MAPE and RMSE values, signifying enhanced accuracy and minimized prediction deviations. However, in certain cases, the ANN model displays a lower MBE in comparison to the LSTM model, indicating reduced bias in its predictions. Regarding the performance of the LSTM model, it achieves a minimal MAPE of 2.8128% for ASTL, signifying precise forecasts with a mere 2.8128% average relative deviation from the true values. For Meezan Bank Limited (MEZL), the Mean Absolute Percentage Error (MAPE) rises slightly to 3.1275%, indicating a somewhat larger average relative difference between the predicted and actual stock prices. In contrast, United Bank Limited (UBL) demonstrates the best performance, with the lowest MAPE among all the companies at just 2.2192%. This reflects the LSTM model's outstanding accuracy in forecasting UBL's closing prices. Likewise, Oil and Gas Development Company (OGDC) shows strong predictive results, with a low MAPE of 3.0984%, further emphasizing the model's ability to deliver precise forecasts with minimal average error.

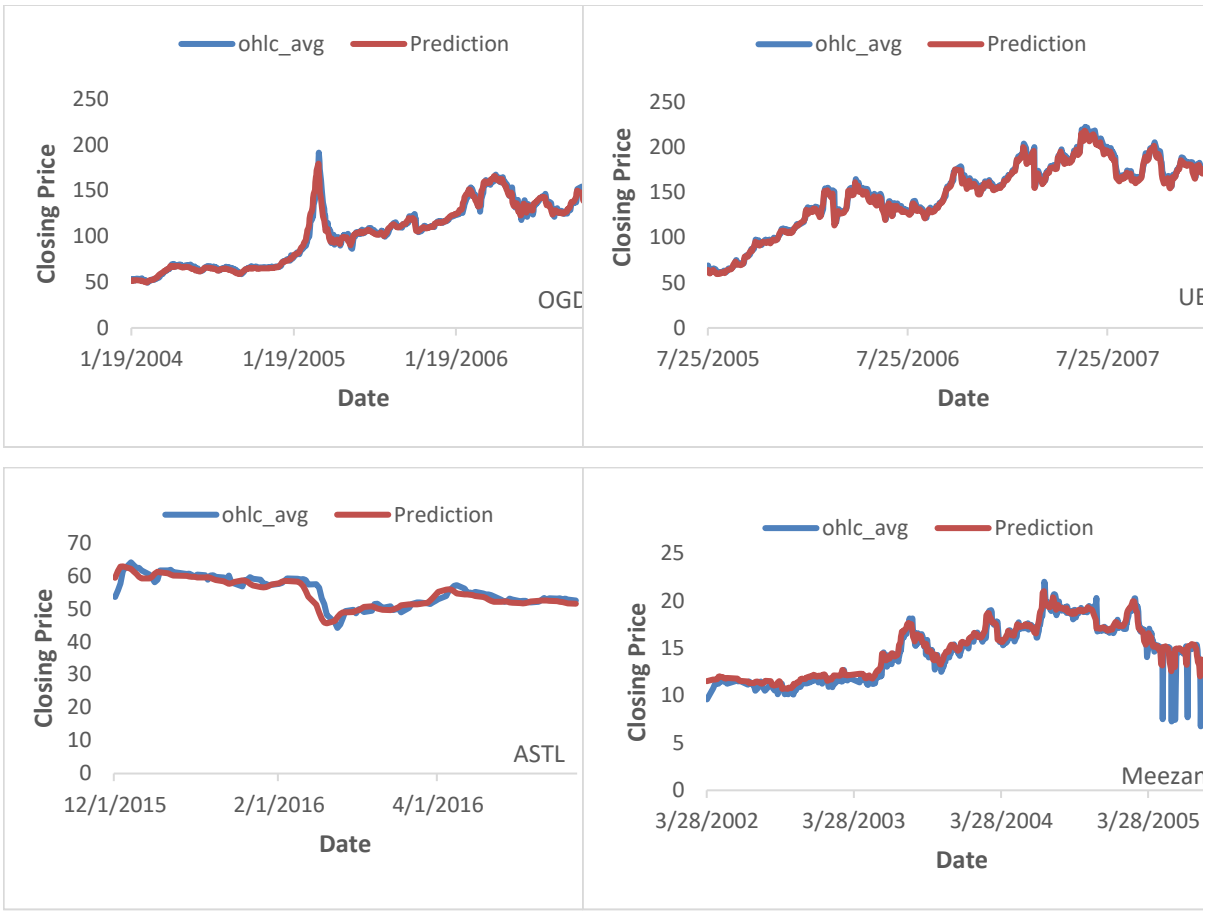

Figure 3: Predicted Price vs Average OHLC of stock exchange using LSTM

As shown in Figure 3, this study applies to the LSTM model to forecast the closing prices of four publicly listed companies—Oil and Gas Development Company (OGDC), United Bank Limited (UBL), Meezan Bank Limited (MEEZAN), and Amreli Steel Limited (ASTL). The prediction process begins with collecting historical OHLC (Open, High, Low, Close) data for each company. From this, the average OHLC value is calculated for each trading day. These averages are then used as input features, while the actual closing prices serve as target values for training the LSTM model. The data is carefully structured and reshaped to suit the LSTM model's input requirements. Once the neural network is configured and trained on this formatted dataset, it is used to predict future closing prices based on new OHLC average values. The results of these predictions are illustrated using line charts, where the x-axis represents the timeline (dates) and the y-axis displays the closing prices. These visual comparisons between actual and predicted prices offer a clearer understanding of the model's performance and reveal how changes in OHLC averages relate to stock price movements for each company. In terms of performance, the LSTM model generally delivers higher accuracy compared to the ANN model. For most of the companies in this study, the ANN model exhibits

slightly higher MAPE (Mean Absolute Percentage Error) values, indicating larger average prediction errors. However, it is worth noting that the ANN model demonstrates lower MBE (Mean Bias Error) for companies like ASTL and OGDC. This suggests that while ANN may be less precise overall, it tends to produce predictions with less systematic bias in some cases. In conclusion, based on the evaluation metrics used, the LSTM model typically outperforms the ANN model by delivering more accurate and reliable stock price forecasts. Nonetheless, the ANN model's reduced bias for certain companies indicates it may still have value depending on the context. It's important to recognize that these findings are specific to the dataset used in this study and may vary with different data or under different market conditions.

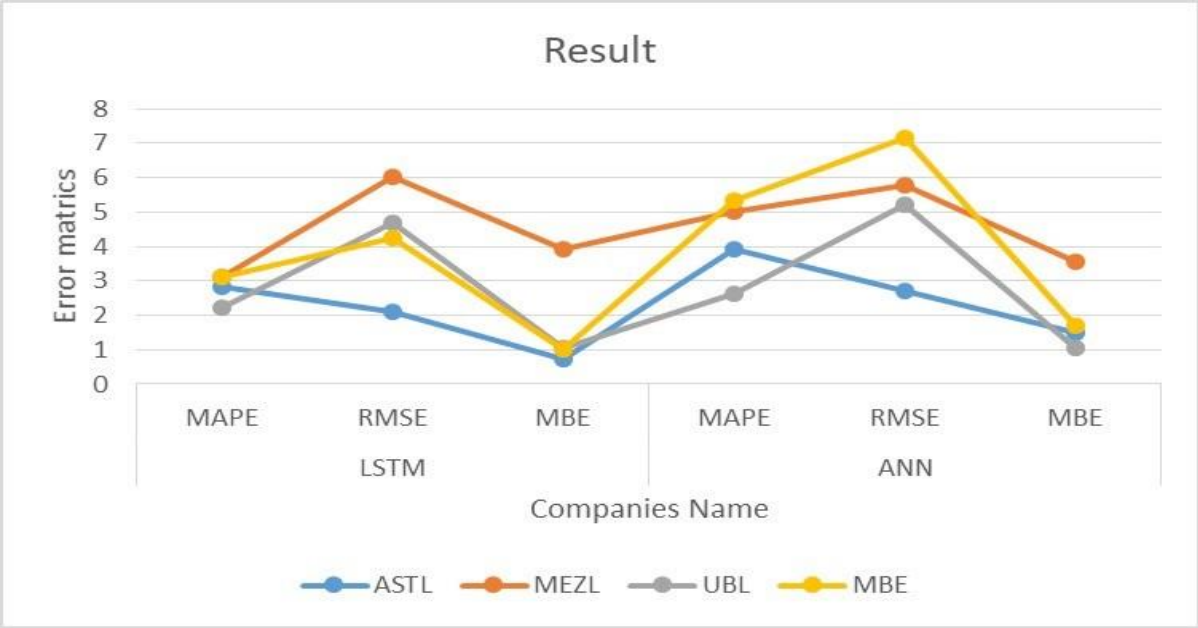

Figure 4 Result Visualization

Figure 4 presents a line chart comparing the performance of LSTM and ANN models across four companies. Overall, the LSTM model demonstrates superior accuracy, reflected in its consistently lower MAPE (Mean Absolute Percentage Error) and RMSE (Root Mean Square Error) values compared to the ANN model. This suggests that LSTM is more effective at closely matching the actual stock closing prices. However, when it comes to bias—measured using Mean Bias Error (MBE)—the ANN model performs better for specific companies, showing a lower tendency to systematically over- or under-predict. For companies like ASTL and UBL, LSTM clearly outperforms ANN in terms of accuracy. In the case of MEZL, the results are mixed, with neither model showing a definitive edge. For OGDC, LSTM again achieves higher accuracy, although it does exhibit a slight positive bias. In summary, LSTM proves to be the more accurate model overall, making it a strong choice for predicting stock

prices based on historical data. That said, the ANN model's lower bias in certain cases suggests it may offer more balanced predictions, which could be beneficial in specific scenarios. A more detailed analysis, considering broader market context and additional variables, would further strengthen the evaluation of both models' performance.

Conclusion & Future Work

Predicting stock closing prices is a difficult undertaking due to the data's complicated and dynamic nature, which is dependent on several elements. Forecasting is critical for lowering company risks, and in recent years, machine learning approaches have become more interactive. This article employs two machine learning algorithms to anticipate the closing prices of publicly traded firms, including ASTL, MEZL, OGDC, and UBL. New variables were created from existing ones to improve prediction accuracy. After applying the LSTM and ANN models to the datasets, the anticipated stock closing prices were tested. Performance measures such as MAPE, MBE, and RMSE are used to compare models. As a result, when compared to ANN, our proposed LSTM model proved to be more successful for the dataset. Integrating social media data and government economic strategies can improve stock closing price projections in the future. Furthermore, using financial news as a metric can provide valuable insights for improved performance. Robust data collecting, natural language processing, and powerful machine learning models will be required to properly incorporate these additional data sources.

References

1. Agrawal, M., Shukla, P. K., Nair, R., Nayyar, A., and Masud, M. (2022). Stock prediction based on technical indicators using deep learning model. *Computers, Materials & Continua*, 70(1):287–304.
2. Ali, M. M., Babar, M. I., Hamza, M., Jehanzeb, M., Habib, S., and Khan, M. S. (2019). Industrial financial forecasting using long short-term memory recurrent neural networks. *International Journal of Advanced Computer Science and Applications*.
3. Anand, C. (2021). Comparison of stock price prediction models using pre-trained neural networks. *Journal of Ubiquitous Computing and Communication Technologies (UCCT)*, 3(02):122–134.
4. Anjum, S. (2020). Impact of market anomalies on stock exchange: a comparative study of kse and psx. *Future Business Journal*, 6(1):1–11.

5. Bao, W., Yue, J., and Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, 12(7):e0180944.
6. Boonpeng, S. and Jeatrakul, P. (2016). Decision support system for investing in stock market by using oaa-neural network. In *2016 Eighth International Conference on Advanced Computational Intelligence (ICACI)*, pages 1–6. IEEE.
7. Chen, Y. and Hao, Y. (2017). A feature weighted support vector machine and k-nearest neighbor algorithm for stock market indices prediction. *Expert Systems with Applications*, 80:340–355.
8. Dhakar, D. S. and Shiwani, S. (2023). A survey on machine learning based approaches for stock market prediction. In *Proceedings of the International Conference on Cognitive and Intelligent Computing: ICCIC 2021, Volume 2*, pages 601–616. Springer.
9. Eapen, J., Bein, D., and Verma, A. (2019). Novel deep learning model with cnn and bi-directional lstm for improved stock market index prediction. In *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)*, pages 0264–0270. IEEE.
10. Emioma, C. and Edeki, S. (2021). Stock price prediction using machine learning on least-squares linear regression basis. In *Journal of Physics: Conference Series*, volume 1734, page 012058. IOP Publishing.
11. Gao, P., Zhang, R., and Yang, X. (2020). The application of stock index price prediction with neural network. *Mathematical and Computational Applications*, 25(3):53.
12. Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02):107–116.
13. Hu, Z., Zhao, Y., and Khushi, M. (2021). A survey of forex and stock price prediction using deep learning. *Applied System Innovation*, 4(1):9.
14. Kaggle (202). Data science company. Last Accessed December 30,2020.
15. Khan, F. A. and Ahmad, N. (2017). Determinants of dividend payout: an empirical study of pharmaceutical companies of pakistan stock exchange (psx). *Journal of Financial Studies & Research*, 16:77–93.

16. Khan, W., Ghazanfar, M., Asam, M., Iqbal, A., Ahmad, S., and Khan, J. A. (2016). Predicting trend in stock market exchange using machine learning classifiers. *Science International*, 28(2).
17. Kumar, D., Sarangi, P. K., and Verma, R. (2022). A systematic review of stock market prediction using machine learning and statistical techniques. *Materials Today: Proceedings*, 49:3187–3191.
18. Leng, J., Liu, W., and Guo, Q. (2022). Stock movement prediction model based on gated orthogonal recurrent units. *Intelligent Systems with Applications*, 16:200156.
19. Masoud, N. M. (2013). The impact of stock market performance upon economic growth. *International Journal of Economics and Financial Issues*, 3(4):788.
20. Mintarya, L. N., Halim, J. N., Angie, C., Achmad, S., and Kurniawan, A. (2023). Machine learning approaches in stock market prediction: A systematic literature review. *Procedia Computer Science*, 216:96–102.
21. Oyewola, D. O., Ibrahim, A., Kwanamu, J. A., and Dada, E. G. (2021). A new auditory algorithm in stock market prediction on oil and gas sector in nigerian stock exchange. *Soft computing letters*, 3:100013.
22. Pagolu, V. S., Reddy, K. N., Panda, G., and Majhi, B. (2016). Sentiment analysis of twitter data for predicting stock market movements. In *2016 international conference on signal processing, communication, power and embedded system (SCOPES)*, pages 1345–1350. IEEE.
23. Pokhrel, N. R., Dahal, K. R., Rimal, R., Bhandari, H. N., Khatri, R. K., Rimal, B., and Hahn, W. E. (2022). Predicting nepse index price using deep learning models. *Machine Learning with Applications*, 9:100385.
24. Ali, A., Raza, R. A. ., & Hassan, A. (2020). Supervised Machine Learning Algorithms: Classification and Comparison Based on Experimental Analysis. Pakistan Journal of Social Sciences, 40(1), 479-492. Retrieved from https://pjss.bzu.edu.pk/index.php/pjss/article/view/857
25. Selvamuthu, D., Kumar, V., and Mishra, A. (2019). Indian stock market prediction using artificial neural networks on tick data. *Financial Innovation*, 5(1):1–12.
26. Shah, J., Vaidya, D., and Shah, M. (2022). A comprehensive review on multiple hybrid deep learning approaches for stock prediction. *Intelligent Systems with Applications*, page 200111.

27. Shah, M. U. D., Shah, A., and Khan, S. U. (2017). Herding behavior in the pakistan stock exchange: Some new insights. *Research in International Business and Finance*, 42:865–873.

28. Shen, J. and Shafiq, M. O. (2020). Short-term stock market price trend prediction using a comprehensive deep learning system. *Journal of big Data*, 7(1):1–33.

29. Singh, T., Bhisikar, S., and Kumar, M. (2023). Stock market prediction using ensemble learning and sentimental analysis.

30. In *Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021*, pages 429–441. Springer.

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

32. Toochaei, M. R. and Moeini, F. (2023). Evaluating the performance of ensemble classifiers in stock returns prediction using effective features. *Expert Systems with Applications*, 213:119186.

33. Umer, M., Awais, M., and Muzammul, M. (2019). Stock market prediction using machine learning (ml) algorithms. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 8(4):97–116.

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

35. Werawithayaset, P. and Tritilanunt, S. (2019). Stock closing price prediction using machine learning. In *2019 17th International Conference on ICT and Knowledge Engineering (ICT&KE)*, pages 1–8. IEEE.

36. Xu, Y. and Cohen, S. B. (2018). Stock movement prediction from tweets and historical prices. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1970–1979.