REST Journal on Banking, Accounting and Business

Vol: 4(1), March 2025

REST Publisher; ISSN: 2583 4746

Website: http://restpublisher.com/journals/jbab/

DOI: https://doi.org/10.46632/jbab/4/1/1

AI in Stock Market Forecasting with Reference to Listed Company in NSE

***M. A. Shakila Banu, Jesu Mariya E**

Jamal Mohamed College (Autonomous), Trichy, Tamil Nadu, India.

*Corresponding author: rspazila@gmail.com

Abstract: *This study looks into the connection between the energy variable and stock prices. The AI technology used in this study to forecast the stock market is helpful in predicting the market price. The various tests investigate the interactions between these two variables in the Indian market through various methods, including descriptive statistics, correlation, the ADF test, and artificial neural networks. By understanding the relationship between the energy variable and stock price participation, you can know about the network of interrelated nodes that connects the stock market and the energy variables. Having a neural network and using deep learning for stock price prediction is beneficial.*

Keyword: *Indian market, Artificial Intelligent, Forecasting, Neural Network*

1. INTRODUCTION

Artificial Intelligence (AI) has become a transformative tool in stock market forecasting, particularly concerning the National Stock Exchange (NSE) of India. By leveraging machine learning algorithms and deep learning models, AI can analyse vast datasets to identify intricate patterns and trends that may elude traditional analytical methods. This capability enhances the accuracy of predictions related to stock prices and market movements. For instance, AI models have been developed to predict the next day's opening value of the Nifty-50 index on the NSE, demonstrating the practical applications of AI in this domain. Furthermore, AI's proficiency in processing and interpreting extensive data enables investors to make more informed decisions, thereby optimizing their investment strategies within the NSE framework. As the financial landscape continues to evolve, the integration of AI into stock market forecasting is poised to play an increasingly pivotal role in navigating the complexities of markets like the NSE.

2. REVIEW OF LITERATURE

Muhammad Faheem Ullah; Aysha Sami Latif; Waqar Ahmad; Moazam Sarfraz AI Impact on Asset Pricing in US Stock Market: Bulletin of Management Review; **Vol. 2 No. 1 (2024)** This research examines the impact of artificial intelligence (AI) on the dynamics of asset pricing in the U.S. stock market. Using data spanning from 2010 to 2023, we conduct an in-depth analysis of various financial metrics and indicators of AI adoption to explore the connection between the integration of AI and fluctuations in stock prices. Our results reveal a significant correlation between the implementation of AI and improved pricing efficiency, especially in the high-tech and financial industries. The study employs partial least squares structural equation modelling (PLS-SEM) to analyse the intricate relationships among the factors that contribute to this effect. Findings suggest that AI-driven trading strategies, sentiment analysis, and predictive modelling have profoundly impacted conventional asset pricing models, prompting a reassessment of current financial theories.

Karlo Puh, Marina Bagić Babac Predicting stock market using natural language processing: American Journal of Business ; 6 April 2023 ISSN: 1935-5181 Stock market price prediction has long been a fascinating topic because of its strong relationship to money production. Recent advances in natural language processing (NLP) have opened up new approaches to this challenge. This study seeks to present a cutting-edge natural language system for stock market prediction using language.

Mohammed Alnemer & Abdal muttaleb Al-Sartawi Artificial Intelligence and Stock Trading Decisions : Artificial Intelligence, Internet of Things, and Society 5.0; 09 November 2023 pp 61–69 This paper presented a concise summary of the importance of artificial intelligence in predicting stock market trends. It outlined several established and effective algorithms utilized for this objective. Furthermore, it addressed the limitations and challenges encountered by AI developers, as well as the methodology involved in applying AI for stock market predictions and decision-making.

Mehar Vijh , Deeksha Chandola , VinayAnand Tikkiwal , Arun Kumar Stock Closing Price Prediction using Machine Learning Techniques: Procedia Computer Science; Volume 167, 2020, Pages 599-606 Forecasting stock market returns accurately is quite a difficult endeavour due to the unpredictable and non-linear characteristics of financial stock markets. The emergence of artificial intelligence and enhanced computational power has led to the development of programmed prediction methods that have shown greater efficiency in forecasting stock prices. In this study, artificial neural networks and random forest algorithms have been employed to predict the closing prices for the following day for five companies across various sectors. Financial metrics, including open, high, low, and close stock prices, are utilized to generate new variables that serve as inputs to the model. The performance of the models is assessed using standard evaluation metrics: RMSE and MAPE. The low values for these metrics indicate that the models are effective in forecasting stock closing prices.

Hemil N. Shah prediction of Stock Market Using Artificial Intelligence: 29-31 March 2019 2019 IEEE 5th International Conference for Convergence in Technology (I2CT) 10.1109/I2CT45611.2019.9033776 Across the globe, there are substantial investments in stock markets. The financial systems worldwide are linked to these markets. In contemporary times, trading in stock markets has turned into a means of generating profit. Artificial intelligence is employed to uncover unknown data and analyse it, using these insights to forecast stock market trends. Forecasting the stock market is a highly complex task that employs various methods, including artificial neural networks (ANN), adaptive neuro-fuzzy inference systems (ANFIS), and swarm intelligence, among others. Future techniques like Levenberg-Marquardt, which is a component of neural networks, not only provide accurate predictions but are also more efficient than ANFIS in several aspects such as time efficiency, memory usage, and accuracy.

3. OBJECTIVES OF THE STUDY

- To examine the relationship between the energy variables and the stock market.
- To find out if a time series dataset contains a unit root.
- To examine a network of interrelated nodes that connects the stock market and the energy variables.

Scope of the Study

In order to predict market trends and price movements, the study of artificial intelligence (AI) in stock market forecasting includes an examination of sophisticated machine learning algorithms, deep learning models, and data analytics approaches. The purpose of this study is to compare the efficiency of AI to conventional financial models in processing massive datasets, finding trends, and producing precise projections. The incorporation of AI tools into investing strategies, risk management, and real-time decision-making is also examined. With insights into potential advancements and useful uses in financial markets, the paper also discusses the difficulties, constraints, and moral dilemmas related to AI-driven stock market forecasting.

4. RESEARCH METHODOLOGY

Research Design:

The research study is carried out, and the plan is created for the gathering of historical data on the explanatory variable and taking stock market returns into account for data measurement and analysis among other research designs. A strategy that describes how to carry out a research study is called a research design. It describes your research topic, what you hope to learn, and how you plan to gather and analyse the data. Making sure the study is well-structured and the findings are reliable is the aim. Research designs come in a variety of forms:

- To test concepts or identify trends, quantitative design makes use of facts and numbers.
- The goal of qualitative design is to use observations and interviews to gain insight into experiences, emotions, or behaviours.
- Mixed-methods provides a more comprehensive view by combining both methodologies.

Period and Nature of the Study:

The secondary data associated with the stock and energy markets comprises five years, from 2020 to 2024. The researcher chose this time frame since it marked the start of the new AI era. The researcher exclusively used historical data from the NSE website for the purpose of this study. Secondary data is information that has already been collected by another entity for a different reason. It is previously collected data that is available for use in administrative, commercial, scientific, or research contexts.

5. DATA ANALYSIS

Descriptive Statistics:
Descriptive statistics provides basic information about the data. It comprises Jarque-Bera statistics as well as the mean, median, minimum, maximum, standard deviation, skewness, and kurtosis. This aids in comprehending the characteristics of data distribution.

Mean:
If the mean is measured inside the confidence intervals, it is primarily a useful indicator of the variable's central tendency. The range of values surrounding the mean that, with a certain degree of certainty, we anticipate the genuine (population) mean to be found is provided by the confidence interval for the mean.

$$\text{Mean}= \sum xi\ n\ \Sigma X \qquad (1)$$

i represents the total of all the value , n represents the total no of frequency

Median:
The mid value of a sample, where half of the observations lie above and half lie below the median when ordered in either ascending or descending order, is determined by the median, which is also a valuable indicator of the "central tendency" of the variable. The median, which is the average of the middle values, is determined when the total number of observations is even.

$$\text{Median}=(n+1)/2 \qquad (2)$$

Maximum & Minimum:
Throughout the study time, the maximum value is the value that is found to be highest, and the minimum value is the value that is determined to be lowest. The highest value attained and the lowest value throughout the study period are represented by the maximum and minimum values, respectively.

Skewness:
Asymmetrical or skewed refers to a frequency distribution of the collection of values that is not symmetrical (normal). Extreme values in a data set shift to one side, or the tail, of a skewed distribution, creating a long tail. The distribution is favorably skewed if the extreme value moves toward the right tail and negatively skewed if it moves toward the left tail.

Skewness = 3(Mean- Median)/ standard deviation

Kurtosis:
"Kurtosis measures the peakedness of a distribution. If the kurtosis is different than 0, then the distribution is either flatter or more peaked than normal".

Jarque – Bera:
The Jarque–Bera test is a goodness-of-fit assessment used in statistics to determine if sample data's skewness and kurtosis match those of a normal distribution. The exam bears the names of Anil K. Bera and Carlos Jarque. There is always a nonnegative test statistic. It indicates that the data do not have a normal distribution if it is far from zero.

Probability:
By randomly choosing a small group of individuals (a sample) from a larger population, probability sampling is a sampling technique that involves estimating the likelihood that the total sample's replies will match the population's overall responses.

Correlation:
Since it enables you to comprehend the links between variables and, consequently, develop hypotheses as the next stage of the process, correlation is a crucial component of any research study. Correlational research frequently aims to identify relationships, characterize them, and then formulate predictions. By comparing two data objects' qualities and computing a score between -1 and +1, the statistical technique known as correlation determines how closely they are related. The data have a positive correlation of +1. Zero means there is no correlation. The data have a negative correlation of -1.

Augmented Dickey-Fuller (ADF) Test:
In order to identify whether time series data is stationary or non-stationary and to check for the presence of a unit root, researchers must use the Augmented Dickey-Fuller (ADF) Test. A crucial premise of many statistical models and forecasting techniques is stationarity. The statistical test that determines whether a time series is stationary or has a unit root (i.e., is non-stationary) is called the Augmented Dickey-Fuller Test, or ADF Test. By leveraging lagged differences to handle data autocorrelation, it enhances the fundamental Dickey-Fuller Test. It determines whether data requires transformation or differencing in order to become steady.

Purpose of the ADF Test
- To determine whether a time series has a unit root, which would suggest that the data is not stationary.
- Because many forecasting models, such as ARIMA, require stationary data, stationarity is crucial.

Ensure Data Stationarity
- **Stationary Data:** Autocorrelation, variance, and mean remain constant across time.
- **Non-Stationary Data:** Random walks, seasonality, or trends can skew model results.

How the ADF Test Works
The test estimates the following regression:

$$\Delta Y_t=\alpha+\beta t+\gamma Y_{t-1}+\delta_1 \Delta Y_{t-1}+\cdots+\delta_p \Delta Y_{t-p}+\epsilon t \qquad (3)$$

Where:
- Yt is the first difference of the time series ($Y_t - Y_{t-1}$)
- α is a constant (optional)
- β is a trend component (optional)
- γ is the coefficient tested for the unit root
- δ are lagged difference terms to account for autocorrelation
- ϵt is the error term

Hypotheses of the ADF Test
- **Null Hypothesis (H_0):** The time series has a unit root (non-stationary)
- **Alternative Hypothesis (H_1):** The time series is stationary

Decision Rule
- If the p-value is less than 0.05, reject H0: the data is stationary.
- If the p-value is greater than 0.05, fail to reject H0: the data is non-stationary.

Unit Root Test
A statistical technique known as a unit root test is used to ascertain whether a time series variable has a unit root and is non-stationary. A time series' mean, variance, and covariance are also evaluated for temporal independence.

$y_t = y_{t-1}$ + stationary process (4)

is the equation for the unit root test? To determine if a time series variable is non-stationary and has a unit root, statisticians use the unit root test. Generally speaking, the existence of a unit root is the null hypothesis, and depending on the test, the alternative hypothesis might be either explosive, trend, or stationary.

Artificial Neural Network:
Through pattern recognition and precise prediction, Artificial Neural Network (ANN) analysis helps solve complicated problems. Because it can handle complex and non-linear data without requiring manual adjustments, it is superior to traditional models. In addition to working effectively with text, numbers, and graphics, ANNs automatically extract valuable information from data. They are extensively employed in fields like as medical diagnosis, financial forecasting, speech processing, and picture recognition. Large datasets and data that is noisy or lacking can also be handled well by ANNs. All things considered, ANN analysis facilitates better decision-making, task automation, and more precise outcomes.

$$y=f\,(i=1\textstyle\sum n \ wixi+b) \qquad (5)$$

Where:
- Xi = Input features
- Wi = Weights assigned to inputs
- b = Bias term

- f (·) = Activation function (e.g., Sigmoid, ReLU, Tanh)
- y = Output of the neuron

6. ANALYSIS

TABLE 1. Descriptive statistics

	Crude oil	Hcl	Infosys	Tcs	Tech Mahindra	Wipro
Mean	71.09939	1100.815	1407.885	3269.849	1134.628	433.7938
Median	73.82252	1091.569	1466.040	3331.342	1102.094	418.0848
Maximum	122.8213	1935.353	1973.143	4552.757	1805.111	720.5111
Minimum	-37.36399	0.000000	0.000000	0.000000	0.000000	0.000000
Std. Dev.	20.10724	335.4917	341.1118	651.0872	304.9964	124.2422
Skewness	-0.593674	0.259408	-0.787454	-0.639431	-0.032249	-0.071748
Kurtosis	4.070942	2.936430	3.053906	3.250590	2.647146	2.716229
Jarque-bera	132.5229	14.16142	128.7148	88.02768	6.669182	5.241237
Probability	0.000000	0.000841	0.000000	0.000000	0.035629	0.072758
Observations	1244	1244	1244	1244	1244	1244

TABLE 2. Correlation Hypotheses

There is a Relationship Between the Energy Variable and Stock Market Price

	Crude oil	HCL	Infosys	TCS	Tech mahindra	Wipro
Crude oil	1	0.53	0.78	0.71	0.68	0.71
HCL	0.53	1	0.83	0.94	0.86	0.65
Infosys	0.78	0.83	1	0.93	0.92	0.87
TCS	0.71	0.94	0.93	1	0.91	0.78
Tech mahindra	0.68	0.86	0.92	0.91	1	0.85
Wipro	0.71	0.65	0.87	0.78	0.85	1

TABLE 3. Test to Measure Stationary (Augumented Dickey-Fuller Test)

	Crude Oil		HCL		INFOSYS		TCS		Tech Mahindra		WIPRO	
	T Statics	P Value	T Statics	P Value	T Statics	P Value	T Statics	P Value	T Statics	P Value	T Statics	P Value
1 %	-3.435 169	0.4659	-3.435 385	0.0039	-3.435 394	0.0007	3.435 398	0.0000	-3.435 398	0.0008	-3.435 390	0.0167
5 %	-2.863 556		-2.863 651		-2.863 655		-2.863 657		-2.863 657		-2.863 653	
10 %	-2.567 893		-2.567 944		-2.567 946		-2.567 947		-2.567 947		-2.567 945	

TABLE 4.

Particulars	Crude Oil	HCL	INFOSYS	TCS	Tech Mahindra	WIPRO
R-squared	0.076792	0.011015	0.014020	0.021537	0.013688	0.008508
Adjusted R-squared	0.073940	0.010220	0.013226	0.020749	0.012894	0.007710
S.E. of regression	2.702806	57.21134	59.97976	132.5717	53.97624	14.30064
Sum Squared resid	9460.184	4071783	4468184	21810902	3615572	254203.9
Log likelihood	-3134.693	-6809.250	-6857.104	-7837.435	-6720.500	-5077.657

TABLE 5.

F- statistic	26.92933	17.66011	17.66011	27.31622	17.22290	10.66620
Mean dependent Var	-0.005411	0.591522	0.591522	1.743045	0.612305	0.198151
S.D. dependent var	2.808636	60.38038	60.38038	133.9689	54.32761	14.35609
Akaike info criterion	4.830297	11.02750	11.02750	12.61373	10.81657	8.160091
Schwarz criterion	4.850182	11.03574	11.03574	12.62198	10.82482	8.168327
Hannan – Quinn criter	4.837758	11.03060	11.03060	12.61683	10.81967	8.163188
Durbin–Watson stat	2.004546	1.124896	1.124896	1.084023	1.120162	1.607642
Prob (F-statistic)	0.000000	0.000028	0.000028	0.000000	0.000036	0.001121

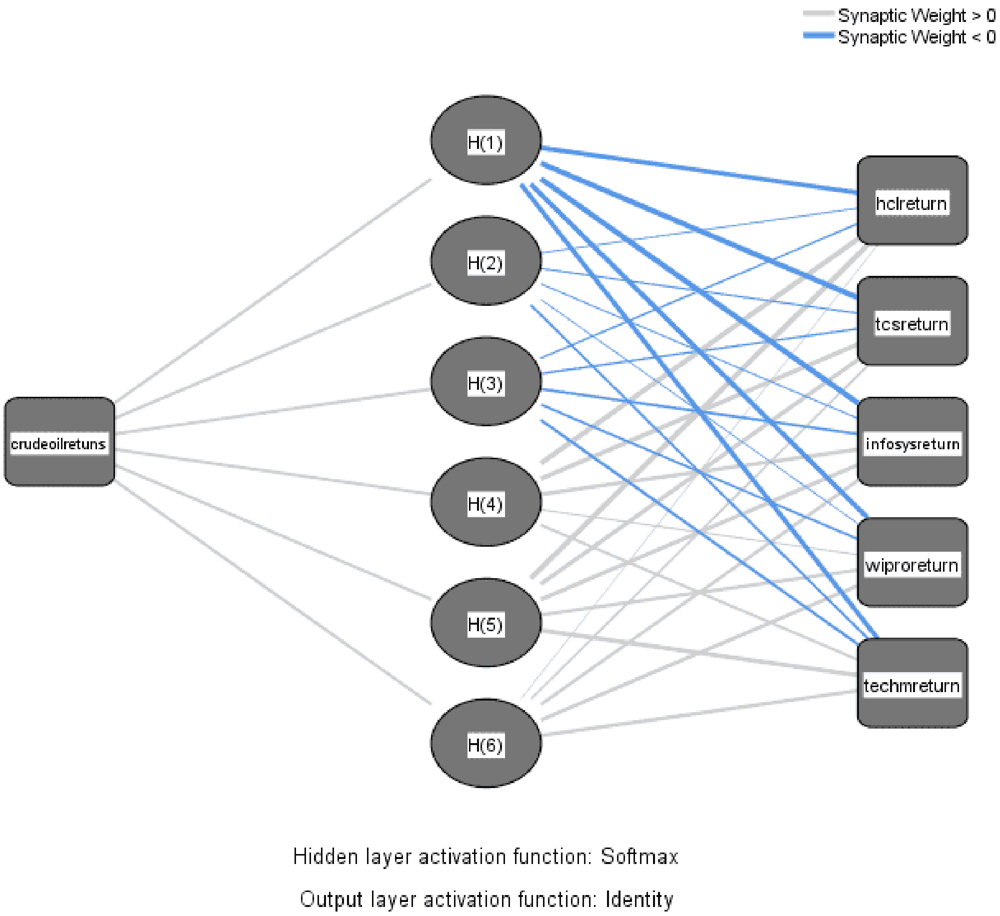

FIGURE 1. Artificial Neural Networks (ANN)

TABLE 6.

Case Processing Summary			
		N	**Percent**
Sample	**Training**	865	69.5%
	Testing	379	30.5%
Valid		1244	100.0%

Network Information			
Input Layer	**Covariates**	**1**	crude oil returns
	Number of Units		1
	Rescaling Method for Covariates		Standardized
Hidden Layer	**Number of Units**		6[a]
	Activation Function		SoftMax
Output Layer	**Dependent Variables**	**1**	HCL return
		2	TCS return
		3	Infosys return
		4	Wipro return
		5	Tech Mahindra return
	Number of Units		5
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares
a. Determined by the testing data criterion: The "best" number of hidden units is the one that yields the smallest error in the testing data.			

TABLE 7.

Model Summary			
Training	**Sum of Squares Error**		749.050
	Average Overall Relative Error		.347
	Relative Error for Scale Dependents	**HCL return**	.446
		TCS return	.269
		INFOSYS return	.222
		WIPRO return	.393
		TECH M return	.404
	Training Time		0:00:00.14
Testing	**Sum of Squares Error**		342.579[a]
	Average Overall Relative Error		.348
	Relative Error for Scale Dependents	**hcl return**	.421
		tcs return	.278
		infosys return	.223
		wipro return	.404
		tech m return	.434
a. The number of hidden units is determined by the testing data criterion: The "best" number of hidden units is the one that yields the smallest error in the testing data.			

7. FINDINGS

- From the above table 1 shows that the statistical measures: Crude Oil, HCL, Infosys, TCS, Tech Mahindra and Wipro. It represents the average value of each variable. The Mean is 3269.849, it is accepted. It gives a sense of the central tendency of the data for median is 3331.342. In the characteristics of the distribution for each variable, including their central tendency, skewness, kurtosis from normality.
- From the above table 2 displays correlation coefficients between pairs of companies. There's a strong positive correlation of 0.78, 0.71 between Crude Oil and Infosys, TCS, Wipro, indicating a tendency for their values to move in the same direction. Crude oil and HCL, Tech Mahindra exhibit a moderate positive correlation of 0.53,0.68, Suggesting some similarity in their performance trends. Overall, these correlations provide insights into the relationships between the company performance.
- From the above table 3, inferred that the ADF test is used to find out the daily crude oil price returns and to determine the existence of the unit root, and it is found that there is an existence of the unit root. The P value is 0.4659 which is greater than the significant value of 0.05 hence, there is evidence to suggest that the time series is non-stationary. Also, the ADF test is used to find the existence of the unit root in NIFTY 50 stock prices, and it is found that there is the existence of the unit root. The P value is 0.0039, 0.0007, 0.0000, 0.0008, 0.0167 which is greater than the significant value of 0.05 hence, there is evidence to suggest that the time series is non-stationary.
- The ANN model attempts to predict the returns of five technology companies using crude oil returns as the sole input. The relative errors indicate varying prediction accuracies across companies, with Infosys having the lowest error 0.22 and HCL the highest 446 during training. The testing errors are similar, suggesting the model generalizes well to new data. However, the significant errors, especially for HCL and Tech Mahindra, imply that crude oil returns may not be a strong stand-alone predictor for these companies' returns.
- By using correlation, we have found the relationship of the variables.
- With the help of ADF tests, we can find out the non-stationarity of the items.
- We learned how to build and test artificial neural networks to predict the leading IT companies.

Suggestions:

By using advanced machine learning algorithms to evaluate enormous volumes of historical and current data, artificial intelligence (AI) holds great promise for stock market predictions. More accurate stock price and trend forecasts are made possible by AI's ability to reveal hidden patterns and linkages that conventional approaches could miss. To assess market sentiment and predict price changes, unstructured data—such as news articles, social media sentiment, and earnings reports—can be processed using methods like deep learning, natural language processing (NLP), and sentiment analysis. By spotting arbitrage opportunities and controlling risk with predictive analytics, AI-powered solutions can help improve trading tactics. Moreover, adaptive trading algorithms that learn from market activity and get better over time can be created using reinforcement learning. Traders and investors can obtain a competitive edge by combining AI with big data and cloud computing to make quicker and better-informed decisions. To guarantee accurate and solid forecasting, it's crucial to handle issues like model interpretability, overfitting, and the influence of market anomalies.

8. CONCLUSION

By offering advanced tools and methods to evaluate intricate and dynamic market data with previously unheard-of accuracy, artificial intelligence has completely transformed stock market forecasting. AI helps investors to efficiently manage risks, make well-informed decisions, and take advantage of market opportunities by utilizing machine learning, deep learning, and data-driven models. Both institutional and individual investors stand to gain from its potential to democratize access to advanced forecasting techniques. However, data quality, model transparency, and the intrinsic volatility of financial markets must all be carefully taken into account for AI to be applied successfully. Even if artificial intelligence (AI) cannot ensure success, its real-time processing and interpretation of massive volumes of data make it a tremendous tool for navigating the difficulties and unknowns of the stock market

REFERENCES

[1]. Muhammad Faheem Ullah; Aysha Sami Latif; Waqar Ahmad; Moazam Sarfraz AI Impact on Asset Pricing in US Stock Market: Bulletin of Management Review; Vol. 2 No. 1 (2024)
[2]. Karlo Puh, Marina Bagić Babac Predicting stock market using natural language processing: American Journal of Business ; 6 April 2023 ISSN: 1935-5181
[3]. Mohammed Alnemer & Abdal muttaleb Al-Sartawi Artificial Intelligence and Stock Trading Decisions : Artificial Intelligence, Internet of Things, and Society 5.0; 09 November 2023 pp 61–69 This paper presented a concise summary of the importance of artificial intelligence
[4]. Mehar Vijh , Deeksha Chandola , VinayAnand Tikkiwal , Arun Kumar Stock Closing Price Prediction using Machine Learning Techniques: Procedia Computer Science; Volume 167, 2020, Pages 599-606
[5]. Hemil N. Shah prediction of Stock Market Using Artificial Intelligence: 29-31 March 2019 2019 IEEE 5th International Conference for Convergence in Technology (I2CT) 10.1109/I2CT45611.2019.9033776 Across the globe, there are substantial investments in stock