**“AI Driven Insights for Stock Marketing Forecasting”**

A Project Report submitted in partial fulfillment of the requirements for the award of the

degree of

## BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

Submitted by

## D. Dheeraj Kumar (VU21CSEN0101004) SMN Rishitha Varma (VU21CSEN0101013)

**Venu Harsha (VU21CSEN0101412) S.Ramesh Naidu (VU21CSEN0101470)**

Under the esteemed guidance of

**Dr. Smruti Rekha Das(Assistant Professor)**

****

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

# GITAM SCHOOL OF TECHNOLOGY

**GITAM (Deemed to be University)**

## VISAKHAPATNAM

**2024**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

# GITAM SCHOOL OF TECHNOLOGY

**GITAM (Deemed to be University)**

****

# DECLARATION

We hereby declare that the project report entitled ‘AI Driven Insights for Stock Marketing Forecasting’ is an original work done in the Department of Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University) submitted in partial fulfillment of the requirements for the award of the degree of B.Tech. in Computer Science and Engineering. The work has not been submitted to any other college or University for the award of any degree or diploma.

Date: 01-11-2024

|  |
| --- |
| **Registration No(s) Name(s) Signature** |
| **VU21CSEN0101004 D. Dheeraj Kumar Dheeraj** |
| **VU21CSEN0101013 SMN Rishitha Varma Rishitha** |
| **VU21CSEN0101412 Venu Harsha Harsha** |
| **VU21CSEN0101470 S.Ramesh Naidu Ramesh** |

## DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**GITAM SCHOOL OF TECHNOLOGY**

**GITAM (Deemed to be University)**

****

# CERTIFICATE

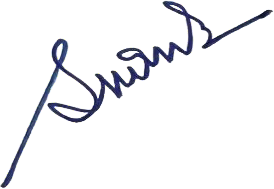
This is to certify that the project report entitled “AI Driven Insights for Stock Marketing Forecasting” is a bonafide record of work carried out by D. Dheeraj Kumar (VU21CSEN0101004), SMN Rishitha Varma (VU21CSEN0101013), Venu Harsha (VU21CSEN0101412), S.Ramesh Naidu (VU21CSEN0101470)

students submitted in partial fulfillment of requirement for the award of degree of Bachelors of Technology in Computer Science and Engineering.

Date : 21-03-2025

Project Guide Head of the Department

Smruthi Rekha Das



# ACKNOWLEDGEMENT

We are truly grateful to everyone who played a role in bringing this capstone project on AI Driven Insights for stock market forecasting to fruition. The journey has been deeply educational, and we owe our sincere thanks to those who offered guidance, support, and resources along the way.

First and foremost, we would like to express our heartfelt appreciation to our project guide and advisor, Dr. Smruti Rekha das, along with reviewers ----. Their consistent support, expert advice, and insightful feedback have been invaluable throughout the research process, helping us to refine and improve this work. We would also like to thank the faculty at GITAM (Deemed to be University), especially the Computer Science Department, for laying a solid groundwork in deep learning which has been crucial in the development of this project.

Finally, we want to acknowledge our colleagues and peers for their constructive feedback, shared insights, and companionship throughout this journey. Their input was instrumental in helping us push the project forward and achieve our objectives.

|  |  |
| --- | --- |
| **TABLE OF CONTENTS** | |
|  | |
| **S.No. Description** | **Page No.** |
| 1. Abstract | 6 |
| 2. Introduction | 6 |
| 3. Literature Review | 6,7,8 |
| 4. Problem Identification & Objectives | 8 |
| 5. Existing System | 8,9 |
| 6. Proposed System Architecture / Methodology | 9,10 |
| 7. Tools/Technologies Used | 10,11 |
| 8. Implementation (Sample Code and Test Cases) | 11.12 |
| 9. Results & Discussion | 12 |
| 10. Conclusion & Future Scope | 13 |
| 11. References | |
| 12. Source Code | |
| 13. Output Screens | |

**1.ABSTRACT**

This research work is based on the collection of Reliance stock market data for supervised learning to train and predict, considering technical indicators such as moving averages, RSI etc. A Hybrid model based on Extreme Learning Machines (ELM), and Artificial Bee Colony (ABC) optimization algorithm is experimented and compared with some of the widely accepted model in stock market. The performance of the proposed model is measured by evaluation metrics such as Mean absolute error (MAE), Mean squared error (MSE), Root mean squared error (RMSE), and R-squared. The research intends to provide a robust predictive model to yield accurate and reliable forecasts through hyperparameter tuning. The results show the prosed model outperform over SLFN and ELM.

## 2.INTRODUCTION

**What is the Stock Market?**

The stock market is a marketplace where investors buy and sell shares of publicly listed companies. It plays a crucial role in the economy by providing a platform for companies to raise capital and investors to trade securities. Stock prices fluctuate based on supply and demand, company performance, and broader economic factors.

**What is the purpose of using AI in stock marketing?**

The purpose of using AI in stock marketing is to make better investment decisions by analyzing large amounts of data quickly and accurately. AI helps predict stock price movements, identify trends, and detect risks based on past market patterns. It automates trading, reduces human errors, and improves decision-making by providing insights that humans might miss. This saves time, increases efficiency, and helps traders and investors make smarter choices in the stock market.

## 3.LITERATURE REVIEW

Previous research has demonstrated the efficacy of AI in stock prediction.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Title** | **Journal/Year** | **Technologies/ Methodologies** | **Result** | **Gap** |
| 1 | “ELM-Based | 2019 | Extreme | Achieved | Application |
|  | AFL–SLFN |  | Learning | better | limited to specific |
|  | Modeling and |  | Machine | performance | industrial |
|  | Multiscale |  | (ELM)Single- | in predicting | processes only |
|  | Model- |  | Layer | tailings grade |  |
|  | Modification |  | Feedforward | in a flotation |  |
|  | Strategy for |  | Neural Network | process. |  |
|  | Online |  | (SLFN) |  |  |
|  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 2 | Stock Price Forecast with Deep Learning | 2021 | LSTM, RNN,  Fully Connected Networks, Convolutional Networks | Best model achieved validation and test MAE of 0.0150 and  0.0148. | Limited to the S&P 500 index; requires testing on other indices or stocks for generalization. |
| 3 | Stock Market Prediction using different Machine Learning Algorithms | 2023 | SVM, Random Forest, LSTM, Decision Trees, Neural Networks | Random Forest achieved the highest accuracy in short-term predictions. | Did not account for external factors like organizational changes, news events, and social media data |
| 4 | Predicting Financial Stock Market Prices Using RNN and LSTM | 2021 | Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM),  Machine Learning | Achieved over 97% accuracy in predicting stock prices using LSTM and RNN. | The prediction accuracy can still be affected by market volatility and requires continuous adjustments for varying datasets. |
| 5 | Emerging Trends in AI- Based Stock Market Prediction: A Comprehensive and Systematic Review | 2023 | AI, Deep Learning, Natural Language Processing, Sentiment Analysis, | Highlights recent advancements and potential implications for investors and policymakers. | Challenges include market unpredictability, data quality, model transparency. |
| 6 | Stock Price Trend Prediction Using | 2023 | Random Forest, Random Search, Machine Learning | The optimized Random Forest showed an average evaluation value that is | The model's performance can still be impacted by data noise and may require continuous |
| 7 | Predicting Financial Stock Market Prices Using LSTM | 2023 | Recurrent Neural Networks (RNN), Long Short-Term Memory | Achieved nearly 97% accuracy in predicting stock prices for selected datasets. | Limited by volatility of the stock market; model performance can vary with different datasets. |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | (LSTM), Deep  Learning |  |  |
| 8 | The Role of Artificial Intelligence Prediction in Stock Market Investors Decisions | 2023 | Long Short- Term Memory (LSTM),  Convolutional Neural Networks (CNN), Deep Learning | The LSTM-  CNN model achieved over 99.98%  prediction accuracy | While the model shows high accuracy, it may be affected by external factors |

**4.PROBLEM IDENTIFICATION AND OBJECTIVES**

Stock market prices keep changing due to various factors like supply and demand, company performance, and economic conditions. Predicting these changes is difficult because:

1. **High Unpredictability** – Stock prices fluctuate frequently, making them hard to predict.
2. **Traditional Methods Fail** – Old prediction techniques are not accurate for dynamic markets.
3. **Large Data Processing** – Analyzing historical stock data manually is time-consuming.
4. **Need for Better Models** – A more efficient AI-based model is required for reliable predictions.

This project aims to use AI to analyze past stock data and find better ways to predict future prices accurately.

**Objectives**

1. **Develop AI Models** – Implement and compare SLFN, ELM, and ELM-ABC models.
2. **Use Technical Indicators** – Apply Moving Averages, RSI, ATR, and statistical measures.
3. **Optimize Accuracy** – Use ABC optimization to improve predictions.
4. **Evaluate Performance** – Measure error rates using MAE, MSE, RMSE, and R².
5. **Provide Reliable Insights** – Help investors make better stock market decisions.

## 5. EXISTING SYSTEM

**Existing System**

**Traditional stock market prediction methods rely on statistical models and basic machine learning techniques. These include:**

1. **Simple Moving Averages (SMA) – Uses past stock prices to find trends but fails to capture sudden changes.**
2. **Linear Regression Models – Predict stock prices based on past data but struggle with market volatility.**
3. **Neural Networks (SLFN) – Performs better but lacks optimization for selecting important features.**

**Drawbacks of the Existing System**

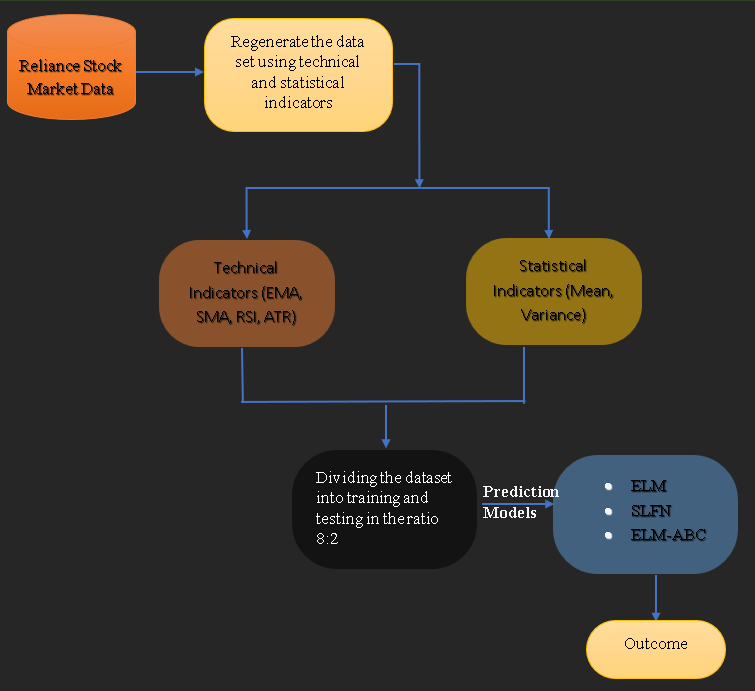
* **Low Accuracy – Struggles to handle rapid price fluctuations.**
* **Limited Feature Selection – Fails to pick the most important factors affecting stock prices.**
* **Slow Processing – Requires large computations, making predictions inefficient.**
* **Inconsistent Predictions – Works well for short-term forecasts but weak for longer periods.**

## 6.PROPOSED SYSTEM ARCHITETURE :

**The proposed system introduces a hybrid AI-driven stock prediction model that combines Extreme Learning Machine (ELM) and Artificial Bee Colony (ABC) optimization.**

**How It Improves Over the Existing System**

1. **Better Accuracy – ELM with ABC optimization improves prediction reliability.**
2. **Feature Selection Optimization – ABC helps filter out less important factors for better performance.**
3. **Faster Processing – ELM allows quick training and execution compared to traditional deep learning models.**
4. **Improved Forecasting for Multiple Timeframes – The model provides reliable short-term (1-day) and multi-day (3, 5, and 7-day) predictions.**



## SYSTEM ARCHITECTURE:

Data Collection

* + Step 1: Collect Stock Market Data

Gather historical data for Reliance Industries, including prices, volumes, and timestamps.

* + Step 2: Apply Technical Indicators

Enhance the data by calculating important indicators:

. True Range (TR):

It shows the stock's volatility on each day by finding the absolute value of the difference between the current day's high and low prices, which is explained in eq (1).

> (1)

Where:

High = The peak price of the share in the ongoing trading session.

Low = The lowest price of the stock in the current trading session.

Open = The opening price of the stock at the beginning of the current trading session.

EMA (Exponential Moving Average): EMA is a type of moving average that places greater weight on recent prices.

Role in AI-driven Prediction: EMA is particularly valuable for detecting short-term trends and reversals in stock prices. By feeding EMA data into machine learning models, we allow the model to adjust its focus on the latest market movements, improving its responsiveness. For example, a model may be able to detect an upward trend earlier by giving more weight to recent price increases.

Formula:

(3)

Where:

Current Price = The most recent price for the stock.

k = The smoothing constant, computed as: k=2/(n+1)

n = no. of periods the moving average is calculated.

Prev. EMA = The EMA value calculated for the previous period.

SMA (Simple Moving Average): SMA is the average of a set number of previous closing prices, calculated by summing them up and dividing by the number of periods.

Role in AI-driven Prediction: SMA is often used to understand the overall direction of the market over longer periods.

Formula:

(4)

Where:

P\_(1,) P\_2………P\_n = The stock price over the last n periods.

n = The amount of periods that are required to find average.

ATR (Average True Range): ATR measures market volatility by calculating the average range between daily highs and lows over a specified period.

Role in AI-driven Prediction: ATR helps machine learning models understand market volatility, allowing them to adjust predictions based on current market conditions.

Formula:

ATR= Average TR over n periods > (2)

Where:

TR = The True Range calculated for every period.

n=The no.of periods over the average is calculated

Statistical Measures:

RSI (Relative Strength Index): RSI is a momentum oscillator that ranges from 0 to 100, measuring the speed and change of recent price movements. Typically, RSI levels above 70 are considered "overbought" (indicating a possible price drop), while levels below 30 are considered "oversold" (indicating a possible price increase).

(5)

Where:

Average Gain = The average of all gains over a timeframe.

Average Loss = The average of all losses over the same period.

Mean and Variance: The mean, or average, of stock prices over a specified period, is a measure of the central tendency of price movements. It provides a general sense of where the stock price typically centers over a period of time.

* + Step 3: Normalize the Data

Scale data for consistency, improving model learning.

* + Step 4: Apply Prediction Models and Compare Accuracy

Test models (SLFN, ELM, and ELM with ABC optimization) to find the one with the best accuracy.

* + Final Outcome: Predicted Stock Prices

Visualize predicted trends alongside actual prices to assess each model’s performance.

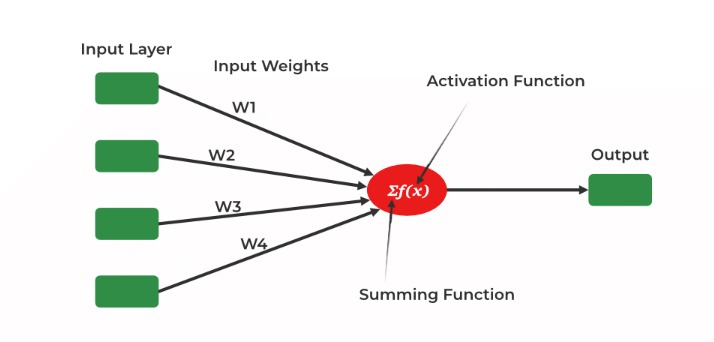
## METHODS TO BE USED :

**Data Preparation:**

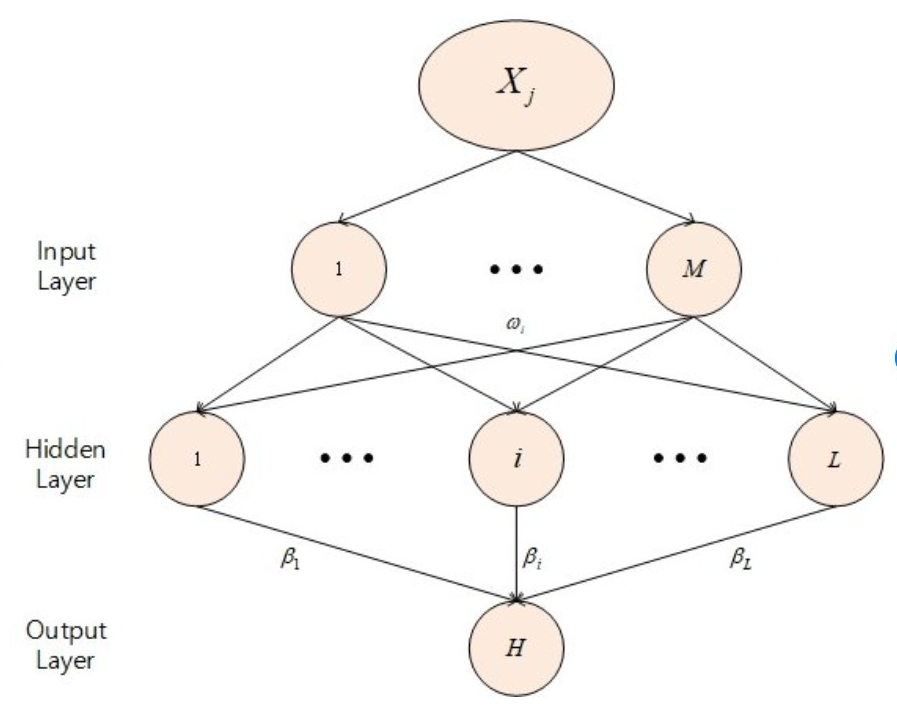
* **Data Cleaning:** Remove or impute missing values, normalize the data, and structure it into a usable format.
* **Excel Formulas:** Moving averages, returns calculation, and volatility measures, to make the data ready for feature selection.

**Prediction Model:**

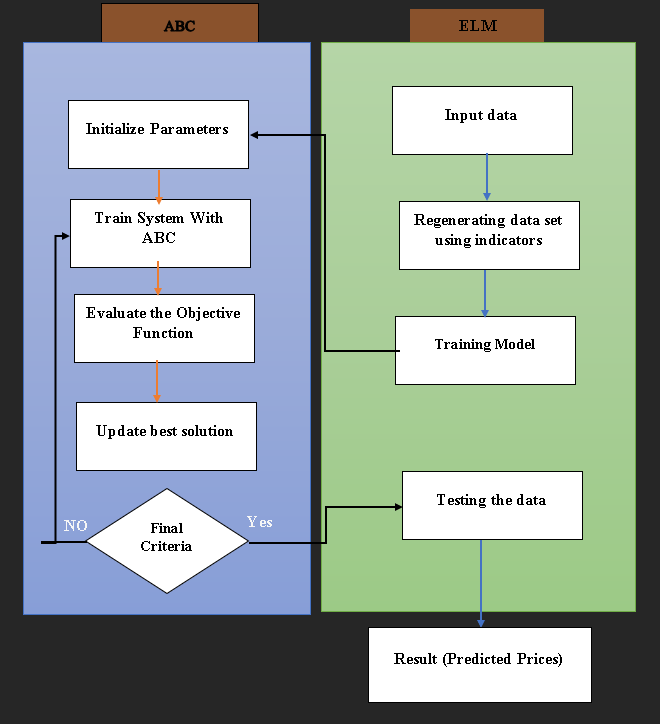
**SLFN:** The SLFN (Single-Layer Feedforward Neural Network) architecture includes only one hidden layer between the input and output layers. Having a single layer makes the SLFN model efficient to train, as fewer parameters are required to optimize compared with deeper neural network architectures.



**Extreme Learning Machine (ELM):** A fast, single-layer neural network used to rank features based on their importance, enabling efficient filtering for the model.



**ELM WITH Artificial Bee Colony (ABC):** A nature-inspired optimization algorithm that filters out unimportant features by simulating the foraging behavior of honey bees. It is used as a learning rule for Neural network model.



**Visualization:** Use graphs to show past stock prices alongside predicted trends, providing a clear view of the model's effectiveness.

## 7.TOOLS AND TECHNOLOGY USED :

1. Data Preparation

Data preparation is the foundational step in any predictive modeling project. In this system, Excel is employed for initial data processing, which includes several key activities:

* + Data Cleaning: This involves identifying and correcting errors or inconsistencies in the data, such as missing values or outliers, ensuring that the dataset is reliable for further analysis.
  + Structuring Data: The cleaned and processed data is organized in a structured format that makes it easy to feed into the subsequent modeling steps. This structured data serves as the input for feature selection, training, and validation.

### Data Collection and Processing:

* **Excel/Google Sheets:** Initial data cleaning and basic calculations (e.g., moving averages).
* **Python:** For data processing and manipulation using libraries like Pandas, NumPy.
* **Yahoo Finance API** : For collecting stock market data.
* Evaluation and Visualization
* **Jupyter Notebook:** For interactive data exploration, model building, and result analysis.

### Data Collection and Processing:

* **Excel/Google Sheets:** Initial data cleaning and basic calculations (e.g., moving averages).
* **Python:** For data processing and manipulation using libraries like Pandas, NumPy.
* **Yahoo Finance API** : For collecting stock market data.
* Evaluation and Visualization
* **Jupyter Notebook:** For interactive data exploration, model building, and result analysis.

## 8. IMPLEMENTATION :

## 8.1 CODING :

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 

## 8.2 TESTING:

## 

## Testing Process Explanation

## Checking Data Quality:

## The code checks if any column in the dataset has only one unique value (univariate column).

## This ensures that all columns contain useful information for predictions.

## Feature Selection:

## The target values (Output\_1, Output\_2, Output\_3, Output\_4) are separated from the dataset.

## The remaining columns are used as input features (X).

## Applying Artificial Bee Colony (ABC) Optimization:

## ABC is used to select the best features from the dataset.

## This helps improve model accuracy by choosing the most relevant data.

## Data Splitting (70% Train, 30% Test):

## The dataset is split into 70% training data and 30% testing data.

## Training data is used to build the model, and testing data is used to evaluate its performance.

## Checking for Univariate Columns:

## Iterates through each column in the dataset.

## If a column has only one unique value, it is considered univariate and might not be useful for predictions.

## Model Training & Evaluation:

## Different regression models (Linear Regression, Ridge, Lasso, ElasticNet) are tested.

## Performance is measured using:

## Mean Squared Error (MSE) → Measures the average error in predictions.

## R² Score → Determines how well the model explains variance in the data.

## The goal is to find the best model with minimal error and maximum accuracy.

## Example Usage of Testing Functions:

## check\_univariate\_columns(stock\_data): Verifies column variability.

## evaluate\_model(trained\_model, X\_test, y\_test): Checks model accuracy on test data.

## 9. RESULT AND DISCUSSION:

## Model Comparison & Performance Analysis

## We tested three models: SLFN, ELM, and ELM-ABC, using Reliance stock data.

## ELM-ABC (Extreme Learning Machine with Artificial Bee Colony Optimization) showed the best accuracy for 1-day predictions, making it ideal for short-term forecasting.

## ELM performed well across multiple days but had slightly higher errors than ELM-ABC.

## SLFN had the highest MSE values, meaning its predictions were less reliable, especially for longer periods.

## Accuracy & Error Analysis

## For 1-day predictions, ELM-ABC had the lowest error, making it the most reliable.

## For 3-day and 5-day predictions, ELM maintained consistent accuracy, but ELM-ABC's performance slightly dropped.

## For 7-day predictions, all models showed reduced accuracy, but ELM was more stable than ELM-ABC.

### 10.CONCLUSION AND FUTURE SCOPE :

The *AI-Driven Insights for Stock Market Forecasting* project demonstrates the effectiveness of AI in predicting stock market trends. In our analysis, we compared three models—SFLN, ELM-ABC and ELM. The results show that the SFLN model provides slightly more accurate predictions overall. However, the ABC model delivers significantly higher accuracy on Day 1, achieving a value of 1, but its performance varies across other days. This comparison highlights the strengths and limitations of each model, offering valuable insights for investors seeking data-driven decision-making.

**Future Scope**

The project lays a foundation for advanced AI-driven stock market analysis. Future improvements may include:

1. **Real-Time Forecasting** – Implementing live stock data updates for real-time predictions.
2. **Deep Learning Integration** – Using **LSTM or Transformers** for enhanced sequential analysis.
3. **Sentiment Analysis** – Incorporating **news and social media data** to refine predictions.
4. **Multi-Stock Prediction** – Expanding the model to predict prices for multiple companies.
5. **Automated Trading** – Developing AI-based **buy/sell decision-making systems** for investors.

* **Accuracy and MSE TABLE:**

## 

## ABC\_ELM shows the lowest MSE and highest accuracy for 1-day predictions — ideal for short-term stock price forecasting.

## ELM maintains decent accuracy across all timeframes, but its MSE is higher than ABC\_ELM for 1-day predictions.

## SLFN performs the weakest overall, especially for longer horizons.

11. REFERNCES:

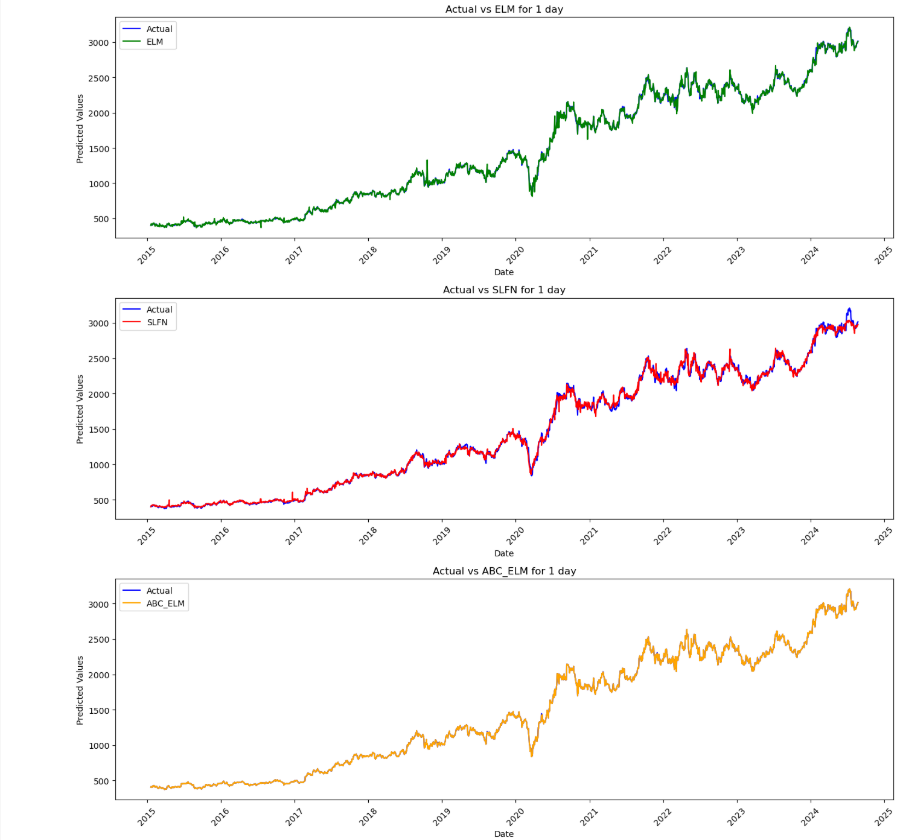
1.O. Mane and S. K. Kandasamy, "Stock Market Prediction using Natural Language Processing—A Survey," 2022. [Online]. Available: <https://arxiv.org/abs/2208.13564>.

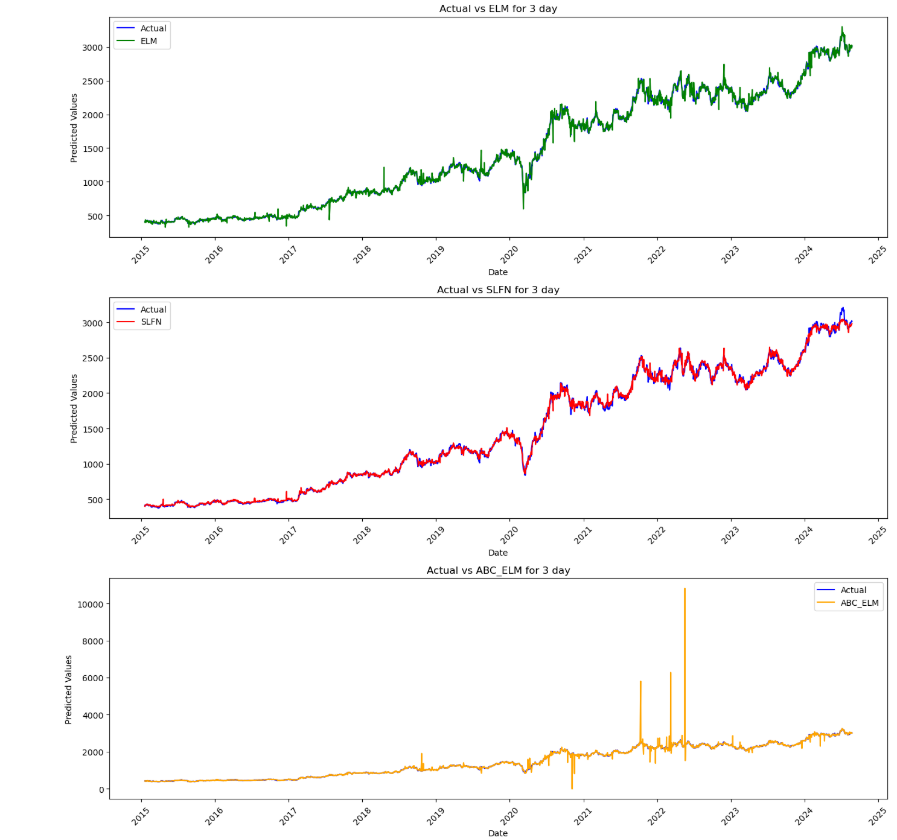
1. S. Mokhtari, K. K. Yen, and J. Liu, "Effectiveness of Artificial Intelligence in Stock Market Prediction based on Machine Learning," 2021. [Online]. Available: <https://arxiv.org/abs/2107.01031>
2. E. Kour, "AI-Driven Predictive Analytics for Financial Market Forecasting," *Int. J. Interdiscip. Finance Insights*, vol. 3, no. 3, 2024. [Online]. Available: <https://injmr.com/index.php/ijifi/article/view/108>

4. S. P. Pattyam, "AI-Driven Financial Market Analysis: Advanced Techniques for Stock Price Prediction, Risk Management, and Automated Trading," *Afr. J. Artif. Intell. Sustain. Dev.*, vol. 1, no. 1, 2021. [Online]. Available: <https://africansciencegroup.com/index.php/AJAISD/article/view/147>

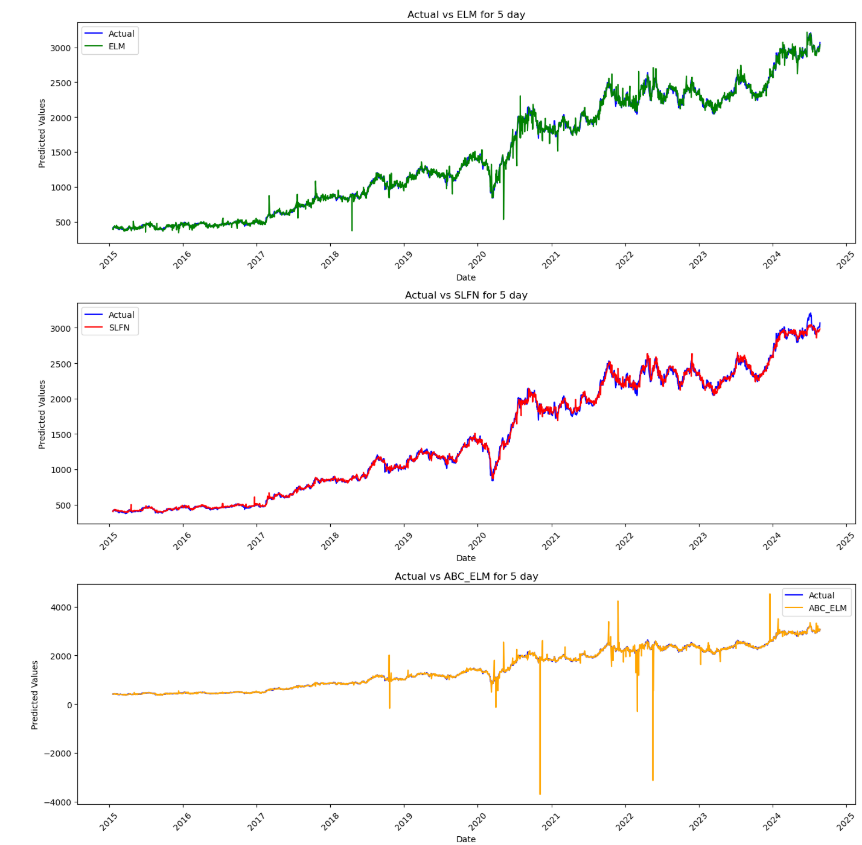
13. Output Screens :

**Comparison of Actual and Predicted Values (Day 1) – Three Models:**

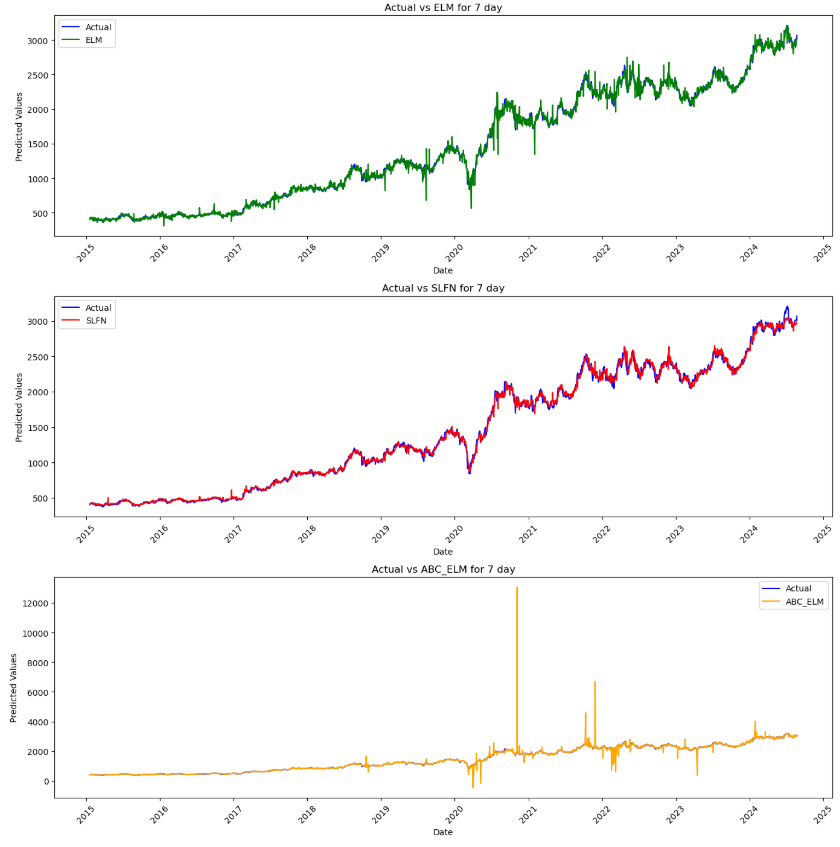
 **Comparison of Actual and Predicted Values (Day 3) – Three Models:**



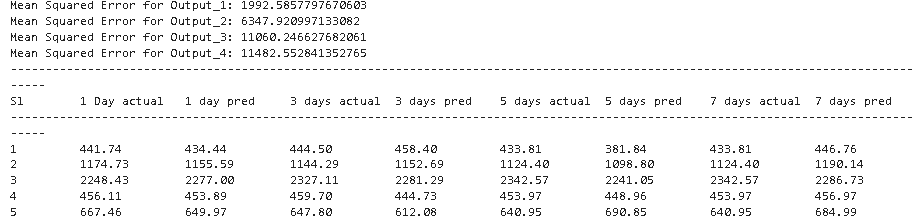
**Comparison of Actual and Predicted Values (Day 5) – Three Models:**



**Comparison of Actual and Predicted Values (Day 7) – Three Models:**

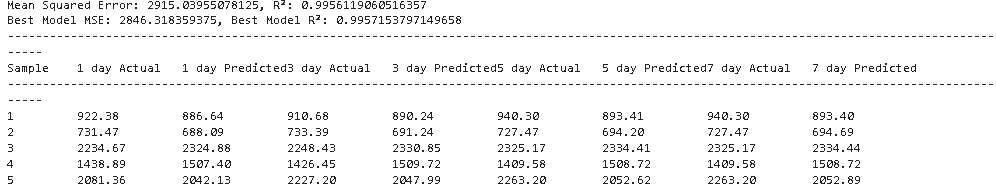


**ELM MODEL:**

****

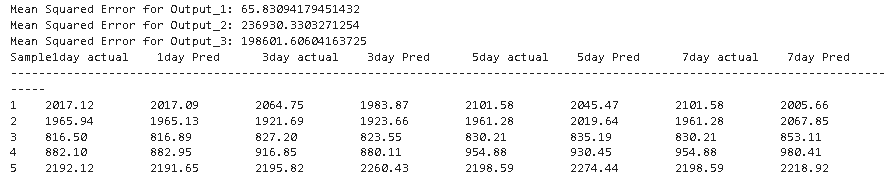
* Shows MSE for multiple outputs, with Output 1 having the lowest error (1992.58).
* Predictions are relatively close to actual values, but there’s noticeable deviation for longer horizons (5-day, 7-day).
* Consistent, but accuracy declines as the forecast window increases

**SLFN MODEL:**

****

* MSE: 2915.04, with R² values very high (~0.995), suggesting strong correlation but slightly worse than ELM for 1-day predictions.
* Predictions track closely, but there’s a visible gap at 7 days.
* The model struggles more than ELM as days increase.

**ELM- ABC MODEL:**

****

* MSE for Output 1 is impressively low (65.83), while Output 2 and 3 have higher errors.
* 1-day predictions are almost perfect, closely matching actual values.
* Accuracy starts to drop sharply for 3-day (0.9290), 5-day (0.9494), and 7-day (0.8888) forecasts.
* ABC\_ELM is best suited for short-term predictions, particularly for 1-day forecasts.
* **Accuracy and MSE TABLE:**

## 

## ABC\_ELM shows the lowest MSE and highest accuracy for 1-day predictions — ideal for short-term stock price forecasting.

## ELM maintains decent accuracy across all timeframes, but its MSE is higher than ABC\_ELM for 1-day predictions.

## SLFN performs the weakest overall, especially for longer horizons.