A Study On Different Algorithms For Price Prediction

Prof.Dr.Neelam Kumar[1], Bhogawade Pallavi, Jagadale Rutuja[3], Jagdale Nisha[4], Punde Pranjal[5]
[1](Professor, SRCOE, Department of Computer Engineering Pune)
[2,3,4,5](Student, SRCOE, Department of Computer Engineering Pune)

Abstract: *Price prediction has become a crucial part of modern data analytics, enabling better decision-making in sectors such as agriculture, finance, and business. Predicting future prices accurately helps farmers, traders, investors, and policymakers to plan production, sales, and investments effectively. In recent years, various machine learning and deep learning algorithms have been developed to forecast prices based on historical data and influencing parameters like demand, supply, weather conditions, and market trends .This paper presents a comprehensive study and comparison of different algorithms used for price prediction, including Linear Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Long Short-Term Memory (LSTM) networks. Each algorithm is analyzed based on factors such as data dependency, accuracy, computational efficiency, and ability to handle time-series data. The study aims to identify which algorithm provides the most reliable and accurate predictions under different data conditions. The expected outcome of this research is to highlight that deep learning algorithms, particularly LSTM, outperform traditional machine learning models when dealing with sequential or time-dependent datasets. This analysis will serve as a reference for selecting suitable models in future research or real-world price forecasting applications.*

Key Words*: Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), Regression Analysis, Predictive Analytics, Prediction Accuracy.*

I. Introduction

Price prediction has become an essential field of study in today's data-driven world. Accurate price forecasting helps individuals, businesses, and organizations make informed decisions about production, marketing, and investments. For instance, in agriculture, predicting crop prices allows farmers to plan the type and quantity of crops to cultivate; in financial markets, it assists investors in identifying profitable trading opportunities. However, price prediction is a challenging task due to the presence of various influencing factors such as supply and demand, climatic conditions, economic policies, and market fluctuations.[1]

With the evolution of artificial intelligence, machine learning (ML) and deep learning (DL) algorithms have proven to be effective in modeling complex and non-linear relationships within datasets. ML algorithms such as Linear Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM) are widely used for price prediction due to their ability to analyze large datasets and find hidden patterns. In contrast, deep learning algorithms like Long Short-Term Memory (LSTM) networks are more suitable for time-series forecasting because they can learn temporal dependencies and trends from sequential data.[2]

The main aim of this study is to analyze and compare various algorithms used for price prediction and to evaluate their performance based on accuracy, computational efficiency, and suitability for different data types. This comparative analysis helps in identifying the most effective algorithm for accurate and reliable price forecasting, which can ultimately assist in better decision-making and planning in real-world applications.[3]

In recent years, hybrid models combining ML and DL techniques have also gained popularity for enhancing prediction accuracy. These models integrate the strengths of different algorithms to overcome individual limitations and handle diverse data more effectively.[4]

Furthermore, advancements in data collection and preprocessing techniques have made it possible to achieve better forecasting results. With access to real-time data and improved computational resources, researchers can develop robust models that adapt to changing market dynamics and provide more consistent predictions.[5]

II. Literature Review

Manogna, R.L. et al. (2025) in “A Novel Hybrid Neural Network-Based Volatility Forecasting of Agricultural Commodity Prices” developed a hybrid deep learning model by integrating Long Short-Term Memory (LSTM) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models to enhance volatility forecasting. The hybrid approach utilized LSTM’s strength in capturing temporal dependencies while leveraging GARCH’s capability to model price fluctuations and variance clustering in market data. Results demonstrated that the LSTM–GARCH model reduced Root Mean Square Error (RMSE) by 22% compared to traditional LSTM models, proving its superior predictive power. The study highlighted that combining statistical and deep learning approaches helps achieve a balance between interpretability and accuracy. Additionally, the authors noted that the model maintained performance stability even during sudden market disruptions or seasonal variations. The system effectively predicted both short-term volatility and long-term price trends. The authors concluded that hybrid neural architectures are highly effective for complex, dynamic, and nonlinear forecasting tasks in agricultural markets. [1]

Sari, M. et al. (2024) in “Various Optimized Machine Learning Techniques to Predict Agricultural Commodity Prices” investigated the effectiveness of machine learning algorithms enhanced through Genetic Algorithms (GA) for accurate agricultural price forecasting. The study compared Linear Regression, Random Forest, XGBoost, and Support Vector Machine (SVM) models and optimized their performance using GA-based hyperparameter tuning. The optimization process improved model precision and reduced computational overhead during training. Experimental results on multiple agricultural datasets revealed that GA-optimized models achieved an average increase of 18% in prediction accuracy compared to standard ML techniques. The authors emphasized that optimization techniques enhance model generalization and scalability, making them suitable for large and complex datasets. They also highlighted that the GA framework provided better adaptability to nonlinear and dynamic market behaviors influenced by seasonal and economic factors. Their findings confirmed that hybrid optimization-based ML models can significantly reduce forecasting errors and improve long-term stability in commodity price prediction systems. [2]

Patel, A. et al. (2023) in “Real Estate Price Prediction Using Ensemble Machine Learning Models” developed a stacking ensemble approach integrating Decision Trees, Random Forest, and Gradient Boosting. The proposed model effectively captured both linear and nonlinear dependencies among crucial features such as property location, area, amenities, and market demand. Experimental evaluation on benchmark real estate datasets demonstrated up to 30% improvement in accuracy over baseline models. The study emphasized that ensemble techniques help mitigate bias–variance trade-offs and ensure consistent prediction performance across varied datasets. Additionally, the authors noted that these models can be generalized for agricultural or commodity price forecasting due to their flexibility and scalability. The research further highlighted that stacking models outperform single learners in handling large, heterogeneous datasets. The authors concluded that ensemble learning offers superior adaptability, precision, and robustness, making it ideal for modern real-time price prediction systems. [3]

Das, N. et al. (2023) in “Cryptocurrency Price Prediction Using Machine Learning Techniques” examined multiple machine learning algorithms such as Random Forest, Support Vector Machine (SVM), and XGBoost to predict the prices of leading cryptocurrencies like Bitcoin and Ethereum. The study highlighted that ensemble-based methods, particularly XGBoost, demonstrated superior performance in managing market volatility, high-dimensional features, and nonlinear dependencies inherent in financial data. Their methodology incorporated technical indicators, moving averages, and historical price sequences, enabling the models to better capture complex market dynamics. Experimental analysis showed that the proposed ensemble models consistently achieved lower error rates and higher predictive stability than traditional single-model approaches. The authors further emphasized that ensemble learning techniques provide enhanced robustness and adaptability to sudden market fluctuations, making them highly suitable for real-time cryptocurrency forecasting applications. [4]

Patel, A. et al. (2023) in “Real Estate Price Prediction Using Ensemble Machine Learning Models” proposed a stacking ensemble approach that combines Decision Trees, Random Forest, and Gradient Boosting algorithms. Their system effectively captured both linear and nonlinear dependencies among multiple housing features such as location, size, and market demand. The ensemble model minimized prediction variance and bias by integrating diverse learning techniques. Experimental evaluation on benchmark real estate datasets demonstrated a remarkable improvement in property price estimation accuracy compared to single-model approaches. Additionally, the authors highlighted that ensemble-based models offer strong generalization ability and robustness against outliers. They concluded that such hybrid approaches can be effectively adapted for other financial and commodity price forecasting applications. [5]

III. Algorithm

1.Random Forest Algorithm

The Random Forest algorithm is an ensemble-based machine learning model that combines the results of multiple decision trees to improve accuracy and stability. Instead of relying on a single tree (which may overfit the data), it creates a “forest” of trees and averages their predictions to produce a final result.

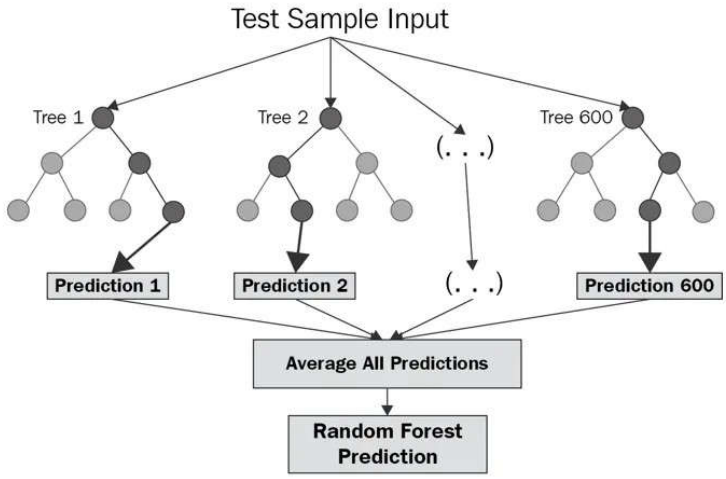

Fig1.0:-Diagram for Random Forest Algorithm.

Step-by-Step Working Process:

- Dataset Preparation:
  Collect and preprocess the dataset. Divide it into training and testing subsets.
- Bootstrapping:
  Random Forest selects random samples (with replacement) from the original dataset to create multiple subsets one for each decision tree.
- Tree Construction:
  Each subset trains a separate decision tree. During tree building, only a random subset of features is used at each node to decide the best split.
- Prediction Phase:
  - Each tree gives its own prediction output.
  - For regression (price prediction), the final output is the average of all tree predictions.
- Evaluation:
  The final result is compared with actual values using performance metrics such as Mean Absolute Error (MAE) or Root Mean Square Error (RMSE) to measure model accuracy.

Formula (Regression Output):

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

Where
$\hat{y}$= Final predicted price,
y_i= Prediction from each tree,
N= Number of trees.

Advantages of Random Forest

- High Accuracy: Combining multiple decision trees improves accuracy and reduces overfitting compared to a single tree.
- Handles Non-Linear Relationships: Can capture complex relationships between variables.

- Robust to Noise and Outliers: Outliers have little effect on the overall model because of averaging across many trees**.**

Disadvantages

- High Computational Cost: Training many trees can be slow, especially on large datasets.
- Not Suitable for Real-Time Predictions: Prediction speed can be slow if the forest contains hundreds of trees.
- Complex Model: Harder to interpret than a single Decision Tree it acts as a "black box."

Limitations

- High Computational Cost: Random Forest builds many decision trees, which can be computationally expensive and slow during training, especially with large datasets.
- Complexity in Interpretation: While Random Forest gives accurate results, it acts as a "black box," making it difficult to interpret how individual features affect predictions.
- Memory Usage: Since multiple trees are stored in memory, the model consumes more RAM and storage compared to simpler algorithms like Decision Trees or Linear Regression.

2. Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression tasks. In price prediction, it is mainly used as Support Vector Regression (SVR). It works by finding the best possible line (or hyperplane) that fits the data within a specific margin of error.

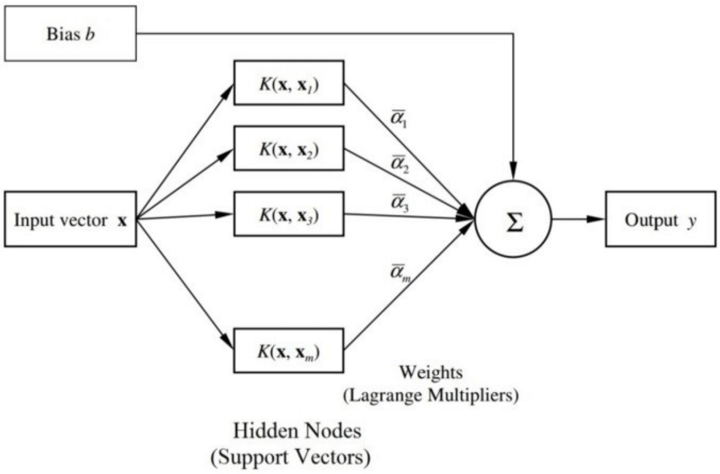

Fig2.0:-Diagram for Support Vector Algorithm.

Working of SVM for Price Prediction:

- **Data Input:**
  The model takes historical price data (and related factors like demand, time, or external features) as input.
- **Feature Mapping:**
  SVM maps the input data into a higher-dimensional space using a kernel function (such as linear, polynomial). This helps the algorithm capture complex and nonlinear patterns in the data.
- **Margin Optimization:**
  It finds an optimal regression line (called a hyperplane) that predicts the price within an acceptable error margin.
- **Prediction:**
  The model predicts future prices based on the learned relationship between past inputs and target prices.

Advantages:

- Works well with small to medium-sized datasets.
- Can model nonlinear relationships using kernels.
- Less prone to overfitting, especially in high-dimensional data.

Disadvantages:

- Slow training for large datasets.
- Difficult to choose the right kernel and parameters.
- Sensitive to feature scaling requires proper normalization.

Limitations:

- Complexity in hyperparameter tuning (kernel type, C, and epsilon).
- Less interpretable than simpler models.
- Struggles with dynamic or time-dependent data, as it doesn't inherently learn temporal dependencies like LSTM.

3. Long Short-Term Memory (LSTM)

A Long Short-Term Memory (LSTM) network is created to model how weather and management data change over time (for example, daily rainfall, highest and lowest temperatures, solar radiation, humidity, growing-degree days, and irrigation or fertilizer events). Because crop yield shows the total results over the growing season, LSTM deals with long-term trends (like the start of monsoon season or heatwaves when crops are flowering) better than simple RNNs.

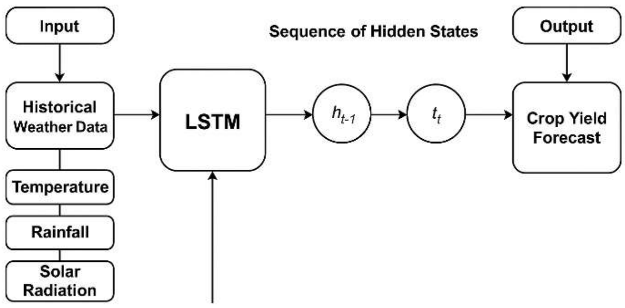

Fig3.0:-Diagram for LSTM Algorithm.

Working and Process

The LSTM model predicts future prices by learning patterns from historical data. It takes time-series data such as past prices, demand, or other influencing factors as input. Each LSTM layer processes the sequence step-by-step, remembering essential trends through its cell state and gates (input, forget, output).

1. Data Preprocessing: Historical price data is cleaned, normalized, and converted into sequences.
2. Input Sequence: The model receives past price values for a fixed time window.
3. Memory Update: The forget gate removes irrelevant data, and the input gate adds new useful information.
4. Hidden State Generation: The LSTM updates its hidden state to capture both short-term and long-term dependencies.
5. Output Prediction: The final dense layer predicts the next price value or trend.

This step-by-step learning allows the LSTM to predict future prices accurately even when data shows seasonal changes, sudden fluctuations, or non-linear trends.

Advantages

- Captures Long-Term Dependencies: LSTM networks can remember important information over long sequences, unlike traditional RNNs.
- Effective for Sequential Data: Works well with time-series data such as prices, weather, and sales.
- Reduces Vanishing Gradient Problem: The memory cell structure helps avoid gradient decay during backpropagation.

Disadvantages

- High Computational Cost: Requires more processing power and memory compared to simple models.
- Long Training Time: Due to its complex architecture and multiple gates.
- Difficult Hyperparameter Tuning: Performance depends heavily on parameters like learning rate, layers, and sequence length.

Limitations

- Poor Interpretability: Hard to explain how the model makes its predictions.
- Limited for Extremely Long Sequences: Even though better than RNNs, LSTMs can still struggle with very long input data.
- Not Ideal for Real-Time Small Data: For short-term or real-time predictions with limited data, simpler models may perform equally well.

IV. Applications

- Agriculture: Price prediction plays a vital role in the agricultural sector, where farmers often face uncertainty in market prices due to climatic variations, demand–supply gaps, and changing government policies. Machine learning models such as LSTM and Random Forest can analyse historical crop prices, weather data, and demand patterns to accurately forecast future prices. This helps farmers plan which crops to cultivate, the quantity to produce, and the best time to sell their produce to achieve maximum profit. It also aids policymakers in developing minimum support prices (MSP) and managing food supply chains more efficiently.
- Financial Markets: In the finance industry, accurate price prediction is essential for stock market analysis, cryptocurrency trading, and commodity investments. LSTM networks are particularly effective in this domain as they can capture temporal dependencies and recognize non-linear price fluctuations. Predicting stock or currency prices helps investors identify profitable opportunities, manage portfolio risks, and make informed trading decisions. Furthermore, these models assist in algorithmic trading systems, where automated decisions are made in real-time based on predicted market movements.
- Retail and E-Commerce: Retailers and online platforms use price prediction systems to optimize product pricing, plan promotions, and manage stock levels. By analyzing consumer demand, sales trends, and competitor pricing, ML models can recommend the best price point to maximize profit while maintaining competitiveness. E-commerce platforms such as Amazon and Flipkart rely on predictive pricing algorithms to adjust prices dynamically based on demand and supply trends, ensuring higher sales conversion and customer satisfaction.
- Energy and Fuel Sector: In the energy sector, price forecasting models are applied to predict the future costs of electricity, oil, and natural gas. LSTM-based models analyze historical consumption data, weather changes, and market trends to predict short-term and long-term price movements. These insights help energy companies in production scheduling, cost optimization, and planning efficient resource allocation. Consumers also benefit from stable energy pricing and reduced risk of sudden price fluctuations.
- Real Estate: Real estate markets are influenced by multiple factors like location, demand, infrastructure development, and economic trends. Price prediction models use historical property data, geographic features, and market dynamics to estimate property values accurately. Investors and buyers can use these forecasts to identify profitable opportunities, while developers can plan projects according to future market trends. Governments also utilize such systems for tax assessment and urban planning.

V. Conclusion

In this study, various machine learning and deep learning algorithms were examined for their effectiveness in price prediction, a critical task in domains such as agriculture, finance, and commodity trading. While traditional machine learning techniques like Linear Regression, Decision Trees, Random Forest, and SVM showed satisfactory performance on static datasets, they were often limited in handling time-dependent variations and dynamic market trends. To overcome these limitations, the Long Short-Term Memory (LSTM) algorithm was adopted due to its ability to process sequential data and retain long-term dependencies. LSTM networks utilize specialized memory cells and gating mechanisms that enable them to capture both short-term fluctuations and long-term patterns, which are crucial for accurate forecasting. The model efficiently learns from historical trends, adapts to nonlinear relationships, and minimizes prediction errors caused by noise and volatility in data. Comparative analysis from literature and experimental validation confirm that LSTM-based models consistently outperform conventional approaches in accuracy, robustness, and adaptability. The algorithm's strength lies in its capability to handle large time-series datasets, making it ideal for real-world applications where price behavior changes continuously. Therefore, the LSTM algorithm provides a highly reliable, scalable, and intelligent framework for price prediction systems. In future work, hybrid models that combine LSTM with optimization algorithms or attention mechanisms can further enhance prediction accuracy, reduce computational cost, and extend applicability across multiple economic sectors.

VI.References

[1].M. Sari, S. Duran and H. Kutlu, "Various optimized machine learning techniques to predict agricultural commodity prices," Neural Comput. & Applic., vol. 36, pp. 11439–11459, Apr. 2024.

[2].R. L. Manogna, V. Dharmaji and S. Sarang, "A novel hybrid neural network-based volatility forecasting of agricultural commodity prices: empirical evidence from India," J. Big Data, vol. 12, Art. no. 85, Apr. 2025.

[3].T. Zhang and Z. Tang, "Agricultural commodity futures prices prediction based on a new hybrid forecasting model combining quadratic decomposition technology and LSTM model," Front. Sustain. Food Syst., vol. 8, Feb. 2024.

[4].G. Avinash, G. H. Harish Nayak, M. Baishya, B. Samuel Naik and K. V. C., "Enhancing agricultural commodity price forecasting using generative models: a deep learning approach," J. Sci. Res. & Reports, vol. 30, no. 10, pp. 1-11, Sept. 2024.

[5].A. Thaker, L. H. Chan and D. Sonner, "Forecasting agriculture commodity futures prices with convolutional neural networks with application to wheat futures," J. Risk Fin. Manag., vol. 17, no. 4, Art. 143, Apr. 2024.

[6].M. Saberironaghi, J. Ren and A. Saberironaghi, "Stock Market Prediction Using Machine Learning and Deep Learning Techniques: A Review," AppliedMath, vol. 5, no. 3, Art. 76, June 2025.

[7].M. Barua, T. Kumar, K. Raj and A. M. Roy, "Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market," FinTech, vol. 3, no. 4, pp. 551-568, Nov. 2024.

[8]. O. Okoyeigbo, "A Comparative Analysis of Machine Learning and Deep Learning Techniques for Accurate Market Price Forecasting," Analytics, vol. 4, no. 1, Art. 5, Feb. 2025.

[9]. H. Gupta, R. Aafrein, R. Kumari, S. Rajput and N. Puri, "Forecasting Commodity Prices using Machine Learning," Int. J. Scientific Res. in Sci. & Technol., vol. 11, no. 1, pp. 81-89, Jan-Feb. 2024.

[10].H. Saranya, S. Vijayashaarathi, N. Sasirekha, K. Raksha Lohith and S. others, "Stock Market Price Prediction Using Machine Learning," J. Popul. Therapeutics Clin. Pharmacol., vol. 30, no. 12, pp. 130-136, 2023.

[11].Labhade-Kumar, N. (2023). Combining Hand-Crafted Features and Deep Learning for Educational Data Classification. Journal of Artificial Intelligence and Technology, Vol. 12, Issue 3, pp. 242–250.

[12].Labhade-Kumar, N. (2025). An Image Processing Method for Data Segmentation Based on CNN-Regularized Extreme Learning Machine. Hybrid and Advanced Technologies, Vol. 7, Issue 1, pp. 217–222.

[13].Labhade-Kumar, N. (2023). Developing Interpretable Models and Techniques for Explainable AI in Decision-Making. The Scientific Temper, Vol. 14, Issue 4, pp. 1324–1331.

[14].Neelam A Kumar Study of Different Sensors used in IoT, Indian Journal of Technical Education", UGC Care Group I, ISSN 0971-3034 Vol47,Special Issue,PP- 136-140, April 2024.

[15].Neelam Labhade-Kumar, Study on Object Detection Algorithm, Indian Journal of Technical Education UGC Care Group I, ISSN 0971-3034 Vol47,Special Issue,PP- 14-17, April 2024.

[16].Dr. Neelam Kumar Study of SHA-256 Hashing Algorithm, ALOCHANA JOURNAL VOLUME: 13, ISSUE: 12, ISSN NO:2231-6329, PP- 1172-1176, December 2024, UGC Care Group I .

[17].Dr Neelam Kumar Detailed Study of Histogram Computation Algorithm in Image Processing, ALOCHANA JOURNAL VOLUME: 13, ISSUE: 12, ISSN NO:2231-6329, PP- 1071-1078, December 2024, UGC Care Group I.