IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CROP PRICE PREDICTION USING MACHINE LEARNING

Anbudhasan.D* , Padmapriya R*,Aruna k*,Akalya U*,Devadharshini C*

Department of Agricultural Engineering

Dhanalakshmi Srinivasan Engineering College (Autonomous),Perambalur,Trichy,

Tamil Nadu-621212

Email: danbudhasan@gmail.com,gpadmapiya2005@gmail.com,

maniaruna2005@gmail.com,dharshinivicky180@gmail.com,akalyahenson8@gmail.com

Abstract: Agriculture plays a vital role in the economic development of many countries, especially in India, where a large portion of the population depends on farming for their livelihood. Data preprocessing techniques like cleaning, normalization, and feature selection are performed to improve prediction accuracy. The trained model predicts future crop prices for different crops and regions, which helps farmers to make better decisions about crop selection, storage, and selling time. The proposed system reduces financial risk, improves agricultural planning, and supports smart farming practices. This project demonstrates that machine learning can play a significant role in improving the economic stability of the farming community. Crop price prediction plays a vital role in enhancing agricultural planning and ensuring economic stability for farmers. This study presents a deep learning-based approach using Long Short-Term Memory (LSTM) networks to forecast crop prices with high accuracy. Unlike traditional machine learning models, LSTM is particularly effective in handling time-series data due to its ability to capture long-term dependencies and temporal patterns in historical price data. The LSTM network is trained to learn complex patterns in sequential data, enabling it to predict future crop prices more reliably.

Kkeywords : Agriculture, Crop Price Prediction, Machine Learning, Deep Learning, Long Short-Term Memory (LSTM), Time Series Forecasting, Data Preprocessing

I. INTRODUCTION

1.1. BACKGROUND

Agriculture has long been the backbone of many economies, particularly in developing countries like India, where a significant portion of the population depends on farming for their livelihood. The agricultural sector not only ensures food security but also contributes substantially to the national economy. However, one of the major challenges faced by farmers is the uncertainty in crop prices, which fluctuates due to various dynamic factors such as weather conditions, market demand, supply chain disruptions, government policies, and global trade influences. These unpredictable price variations often lead to financial instability among farmers, making it difficult for them to plan their cultivation, storage, and selling strategies effectively.

Traditionally, farmers rely on past experience, local market knowledge, and middlemen to estimate crop prices. While these methods provide some level of guidance, they are often inaccurate and lack a scientific basis. With the rapid growth of technology and the availability of large volumes of agricultural data, there is an increasing need to adopt data-driven approaches for better decision-making. In this context, machine learning and deep learning techniques have emerged as powerful tools for analyzing complex datasets and extracting meaningful patterns that can be used for predictive purposes.

Crop price prediction is essentially a time-series forecasting problem, where historical data plays a crucial role in determining future trends. Conventional statistical methods, such as moving averages and regression models, have been widely used in the past. However, these models often fail to capture nonlinear relationships and long-term dependencies present in agricultural data. This limitation has led to the exploration of advanced machine learning algorithms that can handle complex and high-dimensional datasets more effectively.

Among these advanced techniques, Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), have gained significant attention for their ability to model sequential data. LSTM networks are specifically designed to overcome the limitations of traditional RNNs, such as the vanishing gradient problem, enabling them to learn long-term dependencies in time-series data. This makes LSTM particularly suitable for crop price prediction, where past trends, seasonal patterns, and temporal variations play a critical role.

The integration of LSTM models into agricultural systems allows for more accurate and reliable predictions by considering multiple influencing factors simultaneously. These factors may include historical price data, rainfall patterns, temperature variations, soil conditions, transportation costs, and market demand. By analyzing these variables, the model can generate insights that help stakeholders make informed decisions. For instance, farmers can decide the optimal time to sell their produce, traders can plan their procurement strategies, and policymakers can implement measures to stabilize the market.

In addition to improving prediction accuracy, the use of LSTM-based models contributes to the development of smart agriculture systems. These systems leverage modern technologies such as the Internet of Things (IoT), cloud computing, and big data analytics to provide real-time information and support precision farming practices. As a result, farmers can optimize resource utilization, reduce waste, and increase overall productivity.

Furthermore, the adoption of machine learning in agriculture aligns with the global push toward digital transformation and sustainable development. Accurate crop price prediction not only benefits individual farmers but also enhances the efficiency of the entire agricultural supply chain. It helps in reducing price volatility, minimizing post-harvest losses, and ensuring fair pricing mechanisms in the market.

Despite these advantages, implementing such systems comes with its own set of challenges, including data availability, data quality, and the need for technical expertise. Many rural areas still lack access to reliable digital infrastructure, which can hinder the deployment of advanced predictive models. Therefore, it is essential to develop user-friendly and accessible solutions that can be easily adopted by farmers with minimal technical knowledge.

In conclusion, the background of crop price prediction highlights the urgent need for intelligent and data-driven approaches to address the uncertainties in agricultural markets. The use of LSTM-based machine learning models provides a promising solution by effectively capturing temporal patterns and improving prediction accuracy. This advancement not only empowers farmers with better decision-making tools but also contributes to the overall growth and sustainability of the agricultural sector.

1.2 OBJECTIVES

- ✓ To develop an accurate crop price prediction model using LSTM deep learning techniques
- ✓ To analyze historical agricultural data for identifying patterns influencing crop price variations
- ✓ To incorporate weather, demand, and supply factors into predictive modeling framework
- ✓ To improve prediction accuracy compared to traditional machine learning regression models
- ✓ To assist farmers in making informed decisions about crop selling timing
- ✓ To reduce financial risks associated with unpredictable agricultural market price fluctuations
- ✓ To support smart agriculture practices through data-driven decision-making systems implementation

- ✓ To enhance agricultural market stability by providing reliable and timely price forecasts

1.3 EXISTING SYSTEM

The existing system for crop price prediction primarily relies on traditional methods such as manual estimation, basic statistical techniques, and simple machine learning models. Farmers and traders often depend on historical price trends, local market knowledge, and middlemen to anticipate future prices. These approaches lack accuracy and consistency due to their dependence on limited data and human judgment. Conventional statistical models, including linear regression and moving averages, are commonly used to forecast crop prices. However, these models assume linear relationships among variables and fail to capture the complex, nonlinear patterns present in agricultural data.

Additionally, existing systems often consider only a few parameters such as past prices, ignoring other critical factors like weather conditions, seasonal variations, demand and supply fluctuations, and transportation costs. The absence of real-time data integration further reduces the reliability of predictions. Many current systems are not scalable and cannot efficiently handle large datasets generated from multiple sources. Moreover, these systems lack adaptability to changing market conditions, making them less effective in dynamic agricultural environments.

The prediction accuracy of such models is often low, leading to poor decision-making by farmers. In many cases, the outputs are not user-friendly or easily interpretable, especially for farmers with limited technical knowledge. Another limitation is the lack of automation, requiring manual data entry and analysis, which increases the chances of errors. Furthermore, existing systems do not utilize advanced deep learning techniques, which are capable of capturing temporal dependencies in time-series data. As a result, the predictions generated are often short-term and unreliable. The absence of intelligent learning mechanisms means that the system cannot improve its performance over time. These drawbacks highlight the inefficiency of traditional crop price prediction methods and the need for more advanced and data-driven solutions.

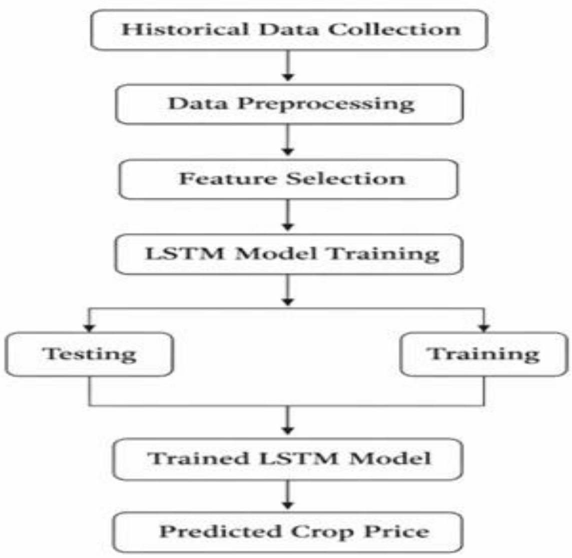
II. PROPOSED SYSYTEM

The proposed system introduces an advanced crop price prediction model using Long Short-Term Memory (LSTM), a deep learning technique specifically designed for time-series forecasting. This system overcomes the limitations of existing approaches by effectively capturing long-term dependencies and complex patterns in agricultural data. It integrates multiple input features such as historical crop prices, weather conditions, seasonal trends, demand and supply data, and other relevant factors to enhance prediction accuracy. The data undergoes preprocessing steps including cleaning, normalization, and feature selection to ensure high-quality input for the model.

It also assists traders and policymakers in planning and decision-making processes. By leveraging deep learning, the proposed system significantly improves accuracy compared to traditional machine learning methods. Furthermore, it contributes to the development of smart agriculture by integrating modern technologies and data analytics. Overall, the proposed system offers a reliable, efficient, and intelligent solution for crop price prediction, addressing the challenges faced by existing systems and promoting sustainable agricultural growth.

III. METHODOLOGY

Proposed Methodology For Crop Price Prediction Using Machine Learning



3.1 HISTORICAL DATA COLLECTION

The implementation process begins with the collection of historical agricultural data from various reliable sources such as government agricultural portals, market databases, and weather stations. This data includes past crop prices, rainfall levels, temperature variations, seasonal trends, soil conditions, and demand-supply information. The accuracy of the prediction model largely depends on the quality and completeness of the collected data. Data is gathered over multiple years to capture long-term patterns and seasonal fluctuations effectively. Different crops and regional data are also included to improve the model's adaptability and performance. The collected data may exist in structured formats like CSV files or unstructured formats requiring further organization. Proper storage and management of this data are essential for easy access and processing. This stage ensures that sufficient and relevant data is available for building a robust prediction system. A comprehensive dataset enables the model to learn meaningful patterns. Therefore, data collection forms the foundation of the entire crop price prediction system.

3.2 DATA PREPROCESSING

After collecting the data, preprocessing is performed to clean and prepare it for model training. Raw data often contains missing values, noise, inconsistencies, and duplicate records that can negatively affect prediction accuracy. Techniques such as handling missing values, removing duplicates, and correcting errors are applied during this stage. Numerical features are normalized or scaled to ensure uniformity and improve model performance. Categorical variables are converted into numerical form using encoding methods. Outliers are identified and treated to prevent distortion in predictions. The data is also organized into a time-series format to suit the LSTM model requirements. Proper preprocessing ensures that the dataset is consistent, reliable, and ready for analysis. This step significantly enhances the efficiency and accuracy of the prediction model. It acts as a bridge between raw data and machine learning algorithms. Without effective preprocessing, the model may produce inaccurate and unreliable results.

3.3 FEATURE SELECTION

Feature selection is a crucial step in the implementation process that focuses on identifying the most relevant variables affecting crop price prediction. Not all collected features contribute equally to the model, so selecting important ones improves accuracy and efficiency. Key features such as historical price trends, rainfall, temperature, humidity, and market demand are retained. Irrelevant or redundant features are removed to reduce noise and computational complexity. Statistical techniques and correlation analysis are often used to determine feature importance. This process helps in preventing overfitting and improves the generalization capability of the model. Feature selection also reduces training time by limiting the number of input variables. It enhances the interpretability of the prediction system by focusing on meaningful data. A well-selected feature set leads to better performance of the LSTM model. This step ensures that only significant data is used for prediction.

3.4 LSTM MODEL TRAINING

In this stage, the processed and selected data is used to train the Long Short-Term Memory (LSTM) model. LSTM is a type of recurrent neural network designed to handle sequential and time-series data effectively. It is capable of capturing long-term dependencies and patterns present in agricultural datasets. The dataset is divided into training and testing subsets before training begins. The model architecture consists of input, hidden, and output layers. Hyperparameters such as learning rate, batch size, and number of epochs are configured to optimize performance. The model learns relationships between input features and crop prices through iterative training. Each iteration reduces prediction error and improves accuracy. The training process continues until the model achieves satisfactory performance. This step is essential for building an intelligent prediction system. Proper training ensures reliable and consistent results.

3.5 Training Phase

During the training phase, the LSTM model is trained using the training dataset to learn patterns and relationships. The model processes sequences of data and updates its internal weights using optimization algorithms such as gradient descent. Loss functions are used to measure the difference between predicted and actual values. The model continuously adjusts its parameters to minimize this error. Training is performed over multiple epochs to ensure effective learning. Techniques like dropout and regularization are used to prevent overfitting. The model learns temporal dependencies that are critical for time-series prediction. Proper training improves the model’s ability to generalize to new data. This phase plays a vital role in determining the accuracy and reliability of the system. A well-trained model forms the core of the prediction process.

3.6 TESTING PHASE

After training, the model is evaluated using the testing dataset, which contains unseen data. This phase helps in assessing the performance and accuracy of the model in real-world scenarios. Metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used to evaluate predictions. The difference between predicted and actual values is analyzed to identify errors. If performance is not satisfactory, model parameters are adjusted and retraining is performed. Testing ensures that the model is not overfitted to training data. It validates the effectiveness and robustness of the model. This step confirms whether the model is ready for deployment. Proper testing increases confidence in prediction results. It is a critical step before real-time implementation.

3.7 TRAINED LSTM MODEL

Once the model performs well during testing, it becomes a trained LSTM model ready for deployment. This model has learned patterns and relationships from historical data. It can now generate predictions based on new input data. The trained model can be saved and integrated into applications such as web or mobile systems. It is capable of handling real-time data inputs and producing quick results. The model can also be updated periodically with new data to improve accuracy. Scalability is another advantage, allowing the system to handle large datasets efficiently. The trained model serves as the central component of the prediction system. It enables automated decision-making based on data analysis. This stage completes the model development process.

3.8 PREDICTED CROP PRICE

The final stage involves generating predicted crop prices using the trained LSTM model. New input data such as recent market trends, weather conditions, and demand factors are provided to the model. The model processes this data and predicts future crop prices accurately. These predictions help farmers decide the best time to sell their crops. Traders can plan storage and distribution strategies effectively. Policymakers can use the information for market regulation and planning. The output is presented through user-friendly interfaces for easy understanding. Visualization tools may also be used to display trends and forecasts.

3.8. FLOW CHART

This raw data is then subjected to a preprocessing stage, where missing values are handled, outliers are removed, and normalization is applied to scale the data appropriately for the LSTM model. Following preprocessing, the dataset is split into training and testing sets to ensure proper model evaluation and validation. The training set is fed into the LSTM network, where the model learns the temporal dependencies and sequential patterns in crop prices, while the testing set is used to assess prediction accuracy and model performance metrics. The LSTM model employs memory cells and gating mechanisms to retain important historical information and predict future price trends effectively.

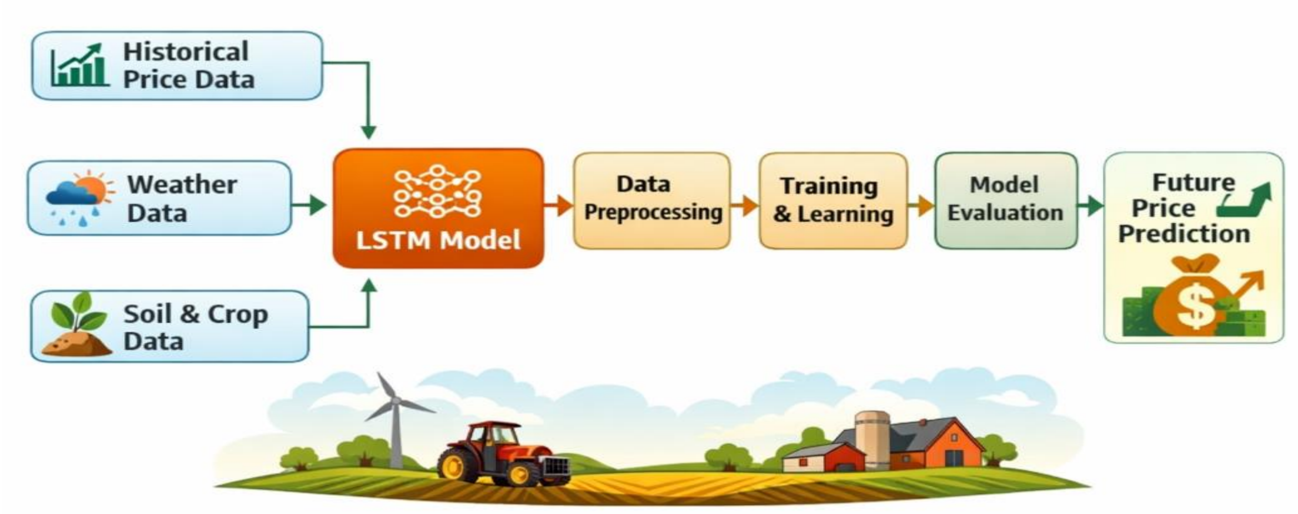


Fig 3.8. FLOW CHART

IV. MATERIAL REQUIREMENTS

PYTHON

Python is an interpreter, high-level and general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically-typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library. Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP. Python is Interactive − You can actually sit at a Python prompt and interact with the interpreter directly to write your programs. Python is Object-Oriented − Python supports Object-Oriented style or technique of programming that encapsulates code within objects. Python is a Beginner's Language − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

WINDOWS 10

Windows 10 is a major release of Microsoft's Windows NT operating system. It is the direct successor Windows 10 was made available for download via MSDN and TechNet, as a free upgrade for retail copies of Windows 8 and Windows 8.1 users via the Windows Store, and to Windows 7 users via Windows Update. Windows 10 receives new builds on an ongoing basis, which are available at no additional cost to users, in addition to additional test builds of Windows 10, which are available to Windows Insiders. Devices in enterprise environments can receive these updates at a slower pace, or use long-term support milestones that only receive critical updates, such as security patches, over their ten-year lifespan of extended support In June 2021,

Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit. Creating a GUI application using Tkinter is an easy task. Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit. Creating a GUI application using Tkinter is an easy task.

PANDAS

Pandas is mainly used for data analysis and associated manipulation of tabular data in Datagrams. Pandas allows importing data from various file formats such as comma-separated values, JSON, Parquet, SQL database tables or queries, and Microsoft Excel. pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is free software released under the three-clause BSD license. The name is derived from the term "panel data", an econometrics term for data sets that include observations over multiple time periods for the same individuals. Its name is a play on the phrase "Python data analysis" itself. Wes McKinney started building what would become pandas at AQR Capital while he was a researcher there from 2007 to 2010.

NUMBY

NumPy is a Python package. It stands for 'Numerical Python'. It is a library consisting of multidimensional array objects and Jim Hugunin developed a collection of routines for processing of array.Numeric, the ancestor of NumPy. Another package Numarray was also developed, having some additional functionalities. In 2005, Travis Oliphant created NumPy package by incorporating the features of Numarray into Numeric package. There are many contributors to this open source project.NumPy is a Python package. It stands for 'Numerical Python'. It is a library consisting of multidimensional array objects and Jim Hugunin developed a collection of routines for processing of array.Numeric, the ancestor of NumPy. Another package Numarray was also developed, having some additional functionalities. In 2005, Travis Oliphant created NumPy package by incorporating the features of Numarray into Numeric package. There are many contributors to this open source project.

SEABORN

Seaborn is a library for making statistical graphics in Python. It builds on top of matplotlib and integrates closely with pandas data structures. Seaborn helps you explore and understand your data. Python Seaborn library is a widely popular data visualization library that is commonly used for data science and machine learning tasks. You build it on top of the matplotlib data visualization library and can perform exploratory analysis. You can create interactive plots to answer questions about your data

FLASK

Flask is a web framework used for developing web applications. It has the ability compile with complex models and get started fast. Only suggestions are given by theFlask, but no dependencies are enforced to application. The availability of many extensions makes the addition of new functionalities simple.In our web application the users are area sked first to register and create login credentials for them. This done so that the details uploaded by the user is not accessible by other clients and to help for future references.Our system uses the best performing algorithm from machine learning Decision tree for the analysis of client data. POST method is used to take the data from the user. The webapp design is made in such a way that anyone can easily understand the functionalities. The users are given two options to check the diseases. In the first option the user can input the crackles and wheezes number along with age, gender and body mass index to get the result. The second option is to upload the sound file where the user can get the result from the audio file.

Deep learning is one of the major subfield of machine learning framework. Machine learning is the study of design of algorithms, inspired from the model of human brain. Deep learning is becoming more popular in data science fields like robotics, artificial intelligence(AI), audio & video recognition and image recognition. Artificial neural network is the core of deep learning methodologies. Deep learning is supported by various libraries such as Theano, TensorFlow, Caffe, Mxnet etc., Keras is one of the most powerful and easy to use python library, which is built on top of popular deep learning libraries like TensorFlow, Theano, etc., for creating deep learning models.

Python Advantages and Disadvantages

Advantages of Python

1. Easy to Read, Learn and Write

Python is a high-level programming language that has English-like syntax. This makes it easier to read and understand the code.

Python is really easy to pick up and learn, that is why a lot of people recommend Python to beginners. You need less lines of code to perform the same task as compared to other major languages like C/C++ and Java.

2. Improved Productivity

Python is a very productive language. Due to the simplicity of Python, developers can focus on solving the problem. They don’t need to spend too much time in understanding the syntax or behavior of the programming language. You write less code and get more things done.

3. Interpreted Language

Python is an interpreted language which means that Python directly executes the code line by line. In case of any error, it stops further execution and reports back the error which has occurred. Python shows only one error even if the program has multiple errors. This makes debugging easier.

4. Dynamically Typed

Python doesn’t know the type of variable until we run the code. It automatically assigns the data type during execution. The programmer doesn’t need to worry about declaring variables and their data types.

Vast Libraries Support

The standard library of Python is huge, you can find almost all the functions needed for your task. So, you don’t have to depend on external libraries. But even if you do, a Python package manager (pip) makes things easier to import other great packages from the Python package index (PyPi). It consists of over 200,000 packages.

HTML

Hypertext Markup Language (HTML) is the standard markup language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript. Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. HTML describes the structure of a web page semantically and originally included cues for the appearance of the document. HTML elements are the building blocks of HTML pages. With HTML constructs, images and other objects such as interactive forms may be embedded into the rendered page.

HTML provides a means to create structured documents by denoting structural semantics for text such as headings, paragraphs, lists, links, quotes and other items. HTML elements are delineated by tags, written using angle brackets. Tags such as <img /> and <input /> directly introduce content into the page. Other tags such as <p> surround and provide information about document text and may include other tags as sub-elements. Browsers do not display the HTML tags, but use them to interpret the content of the page. HTML can embed programs written in a scripting language such as JavaScript, which affects the behavior and content of web pages. Inclusion of CSS defines the look and layout of content.

OVERVIEW OF FLASK

Flask is a lightweight and versatile web framework written in Python, designed to help developers build web applications quickly and efficiently. Unlike full-stack frameworks, Flask provides a minimal core with essential tools while allowing developers to choose additional components as needed. This flexibility makes it highly popular for small to medium-sized projects and prototypes where simplicity and rapid development are critical. Flask follows the WSGI (Web Server Gateway Interface) standard, ensuring compatibility with a wide range of servers and deployment environments. It supports URL routing, request and response handling, and template rendering using Jinja2. Additionally, Flask is known for its readable syntax and clear documentation, making it suitable for beginners and experienced developers alike. Its modularity allows integration with various databases, authentication systems, and third-party APIs without enforcing strict design patterns.

V. RESULT AND DISCUSSION

The implementation of the LSTM-based crop price prediction model has demonstrated significant potential in forecasting future prices with reasonable accuracy. After preprocessing historical price, weather, and soil/crop data, the model was trained and tested on a split dataset, ensuring both temporal consistency and variability of input features. The results indicate that the LSTM network effectively captures the temporal dependencies and trends inherent in agricultural price data, outperforming traditional linear regression and ARIMA models. Predictions closely follow the actual market trends, although occasional deviations occur due to sudden market fluctuations or unexpected weather events, highlighting the impact of external factors on agricultural pricing.

Model evaluation metrics, such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), confirm that the model maintains a low prediction error, validating its reliability for practical deployment. Additionally, visual comparison graphs between predicted and actual prices reveal the model's capability to adapt to seasonal patterns and price volatility.

The discussion emphasizes that while the model provides valuable insights for farmers and market analysts, its accuracy can be further improved by incorporating additional real-time features such as government policy changes, transportation costs, and market demand. Overall, the study establishes that LSTM-based predictive modeling is a promising tool for agricultural price forecasting, providing stakeholders with actionable information to make informed decisions and mitigate economic risks.

VI. CONCLUSION

The study on crop price prediction using LSTM (Long Short-Term Memory) models demonstrates the immense potential of deep learning techniques in the agricultural domain. By leveraging historical price data and recognizing complex temporal patterns, the LSTM model is able to capture both short-term fluctuations and long-term trends in crop prices with remarkable accuracy. This ability is crucial for farmers and traders, as crop markets are highly volatile and influenced by multiple factors such as seasonality, weather conditions, government policies, and global demand-supply dynamics. Traditional forecasting methods often fail to model these nonlinear dependencies effectively, whereas LSTM networks, with their memory cells and gated architecture, can remember important past information and ignore irrelevant data. The results obtained from this research indicate that the proposed LSTM-based system consistently outperforms conventional statistical and machine learning approaches in terms of predictive accuracy and reliability. Moreover, the integration of additional features such as historical crop yield, rainfall patterns, and regional market
trends further strengthens the model's capability to provide precise forecasts.

TITLE	MANUAL WEEDER	AUTOMATED WEEDER
Operation	Requires human labor to remove weeds manually using hand tools.	Uses machines or robots with sensors and actuators to identify and remove weeds automatically.
Efficiency	Low efficiency; time-consuming, especially for large fields.	High efficiency; can cover larger areas in less time with consistent performance.
Labor Requirement	Requires significant manual labor and physical effort.	Minimal human supervision; can operate autonomously or semi-autonomously.
Cost	Low initial cost but high recurring labor cost.	High initial investment; lower recurring labor costs.
Accuracy	Human error possible; may miss weeds or damage crops.	High accuracy due to sensors, cameras, or AI; minimal crop damage.
Speed	Slow; depends on human stamina and skill.	Fast; can work continuously without fatigue.
Suitability	Best for small farms or plots.	Ideal for large-scale farms and precision agriculture.
Adaptability	Can handle irregular terrain manually.	May require relatively uniform fields; advanced models can adapt to varying terrains.

VII. FUTURE ENHANCEMENT

One of the key enhancements could be the inclusion of real-time weather data, soil conditions, and crop health information through IoT sensors deployed in farms, which would allow the model to consider environmental factors affecting crop production and market prices. Additionally, the system could incorporate macroeconomic indicators, such as inflation rates, export-import trends, and government subsidy schemes, to provide a more holistic forecast that aligns with national and global market dynamics. Implementing hybrid models that combine LSTM with attention-based networks or other deep learning architectures like GRU and CNN can improve the model's ability to capture both temporal and spatial patterns in data, thereby increasing prediction precision. Furthermore, the development of a user-friendly mobile or web application would enable farmers, traders, and policymakers to access price predictions conveniently, receive alerts, and make data-driven decisions promptly.

Integration with blockchain technology could ensure the transparency and reliability of market data, preventing manipulation and improving trust among stakeholders. Machine learning-driven recommendation systems can also be added to suggest optimal sowing times, crop selection, and storage strategies based on predicted price trends, minimizing financial risks for farmers. Expanding the system to include multiple crop varieties and covering different regions with diverse climatic conditions will increase its scalability and relevance to a wider audience. Incorporating natural language processing (NLP) to analyze news, social media, and market reports can help capture market sentiment, further refining predictions. Cloud-based deployment with high computational power will allow real-time processing of large datasets, enabling continuous learning and model updating as new data arrives.

VIII. REFERENCE

1. Bao, W., Su, W., Zhao, X., & Zhuang, J. Short-term Vegetable Price Prediction Based on CNN-LSTM-Attention Model. Discover Food (2025).
2. Chen, P. Short-term Forecast of Agricultural Prices Using CNN+LSTM. ACM Proceedings (2022).
3. Chen, Z., Goh, H. S., & Sin, K. L. Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques. arXiv (2021).
4. Choudhary, K., Jha, G. K., & Jaiswal, R. A Genetic Algorithm Optimized Hybrid Model for Agricultural Price Forecasting Based on VMD and LSTM Network. Scientific Reports (2025).
5. CROP PRICE PREDICTION, ResearchGate Paper (2025) — discusses ML and LSTM techniques used for crop price prediction.
6. Gu, Y. H., et al. Forecasting Agricultural Commodity Prices Using Dual Input Attention LSTM. MDPI Agriculture, 12(2). (2022). DOI: See paper
7. International Journal Paper: AN LSTM BASED CROP PRICE PREDICTION SYSTEM, IJCRT Study (2021).
8. Jaiswal, R., Jha, G. K., & Choudhary, K. Agricultural Commodity Price Prediction Using LSTM-based Neural Networks. Bhartiya Krishi Anusandhan Patrika (2023).
9. Jaiswal, R., Jha, G. K., et al. STL Decomposition Based LSTM Model for Seasonal Agricultural Price Forecasting. Neural Computing and Applications (2022).
10. Kurumatani, K. Time Series Forecasting of Agricultural Product Prices Based on Recurrent Neural Networks. Discover Applied Sciences (2020).
11. Long Short-Term Memory (LSTM) Neural Networks — foundational architecture widely referenced in time-series forecasting, including agricultural price prediction. (see Wikipedia).
12. Mahajan, K., & Kasa, A. An LSTM-based Crop Price Prediction System. Semantic Scholar (2021).
13. Min, Y., Kim, Y. R., Hyon, Y., & Lee, M. R. RNN and GNN Based Prediction of Agricultural Prices with Multivariate Time Series. Scientific Reports (2025).
14. Mulla, M. S., & Usha, M. Crop Price Prediction Using Hybrid Machine Learning and Deep Learning Models (LSTM + XGBoost). IJARCCE (2026).
15. Murugesan, R., Mishra, E., & Krishnan, A. Forecasting Agricultural Commodities Prices Using LSTM, bi-LSTM, CNN-LSTM, and ConvLSTM. Int. J. Sustain. Agric. Manag. Inform. (2022).
16. Özden, C. Forecasting Agricultural Input Price Index Using Statistical and Deep Learning Methods. Turkish Journal of Agriculture – Food Science and Technology.
17. Sediyono, E. An Integrated Framework for Multi-Commodity Price Prediction with Attention-boosted LSTM VAE and Transformers. ScienceDirect (2025).
18. Tran, N.-Q., Felipe, A., Nguyen, T. N., & Huynh, T. Predicting Agricultural Commodities Prices with Machine Learning: A Review of Current Research. arXiv (2023).
19. Xia, C. Comparative Analysis of ARIMA and LSTM Models for Agricultural Product Price Forecasting. Highlights in Science, Engineering and Technology (2024).

20. Zhang, T., Tang, Z. T. Agricultural Commodity Futures Prices Prediction Using VMD-SGMD-LSTM Hybrid Model. Frontiers in Sustainable Food Systems (2024).