A Mini Project Report on

Crop Price Prediction and Forecasting using Deep Learning

Submitted in partial fulfilment of the requirements for the award of degree

BACHELOR OF ENGINEERING

in

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

by

CHAKALI PRADEEP (160122729029)

GAJULAPATI ESHWAK(160122729037)

SUNKARI MALLIKARJUN(160122729057)

Under the guidance of

SMT G.KAVITA
Assistant Professor, AI&ML

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY
(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,
Accredited by NAAC with A++ Grade and Programs Accredited by NBA)
Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana
www.cbit.ac.in

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY
An Autonomous Institute | Affiliated to Osmania University
Kokapet Village, Gandipet Mandal, Hyderabad, Telangana-500075, www.cbit.ac.in

COMMITTED TO RESEARCH, INNOVATION AND EDUCATION 46 years

CERTIFICATE

This is to certify that the project titled "**Crop Price Prediction and Forecasting using Deep Learning**" is the bonafide work carried out by students **Chakali Pradeep (160122729029) , Gajulapati Eshwak (160122729037) , Sunkari Mallikarjun (160122729057)** of B.E. AIML of Chaitanya Bharathi Institute of Technology(A), Hyderabad, affiliated to Osmania University, Hyderabad, Telangana(India) during the academic year 2023-2024, submitted in partial fulfilment of the requirements for the award of the degree in **Bachelor of Engineering in Artificial Intelligence & Machine Learning** and that the project has not formed the basis for the award previously of any other degree, diploma, fellowship or any other similar title.

Supervisor	**Head of the Department**
G. Kavita	**Dr. Y. Rama Devi**
Asst. Professor, AI&ML	**Professor, Head of Dept. of AI&ML**

Project Coordinator

Mr. Panduraju Pagidimalla

Asst. Professor, AI&ML

Place: CBIT-Hyderabad

Date:

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE, Accredited by NAAC with A++ Grade and Programs Accredited by NBA)

Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana

www.cbit.ac.in

Vision of Institute

To be the Centre of Excellence in Technical Education and Research.

Mission of Institute

To address the Emerging needs through Quality Technical Education and Advanced Research.

Quality Policy

CBIT imparts value based Technical Education and Training to meet the requirements of students, Industry, Trade/ Profession, Research and Development Organizations for Self-sustained growth of Society.

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,
Accredited by NAAC with A++ Grade and Programs Accredited by NBA)
Chaitanya Bharathi Post, Gandipet , Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana
www.cbit.ac.in

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Vision of the Department

To produce professionals in artificial intelligence and machine learning through the best possible education, acquire international recognition as a destination, and advance society in exciting and creative ways.

Mission of the Department

• Impart rigorous training to generate knowledge through the state-of-the-art concepts and technologies in Artificial Intelligence and Machine Learning.

• Develop technical proficiency in students through creativity and leadership.

• Encourage lifelong learning, social responsibility, environmental conservation, and professional ethics.

• Establish centres of excellence in leading areas of computer and artificial intelligence disciplines.

Program Educational Objectives (PEOs)

PEO1: Using a solid foundation in mathematical, scientific, engineering, and current computing principles, formulate, analyze, and resolve engineering issues.

PEO2: Apply artificial intelligence theory and concepts to analyze the requirements, realize technical specifications, and design engineering solutions.

PEO3: Through cross-disciplinary projects and a variety of professional activities, demonstrate technical proficiency, AI competency, and foster collaborative learning and a sense of teamwork.

PEO4: Provide graduates with solid knowledge, competence, and soft skills that will enable them to ethically contribute to societal demands and achieve sustainable advancement in emerging computer technologies through lifelong learning.

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,
Accredited by NAAC with A++ Grade and Programs Accredited by NBA)
Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana
www.cbit.ac.in

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Program Outcomes (POs)

PO1.Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals and an engineering specialization for the solution of complex engineering problems

PO2.Problem analysis: Identify, formulate, research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

PO3.Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for public health and safety

PO4.Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5.Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools, including prediction and modelling to complex engineering activities, with an understanding of the limitations.

PO6.The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

PO7.Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8.Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice. Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO9.Cosmmunication: Communicate effectively on complex engineering activities with the engineering community and with the society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO10.Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO11.Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Program Specific Outcomes (PSOs)

PSO1. Apply the principal concepts of AI Engineering to design, develop, deploy and prototype AI Subsystems.

PSO2. Apply the knowledge gained pertaining to data storage, data analytics and AI concepts to solve real world business problems.

PSO3. Apply, analyze, design, develop, and test principles of AI concepts on Intelligent Systems.

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

(An Autonomous Institution, Affiliated to Osmania University, Approved by AICTE,
Accredited by NAAC with A++ Grade and Programs Accredited by NBA)
Chaitanya Bharathi Post, Gandipet, Kokapet (Vill.), Hyderabad, Ranga Reddy - 500 075, Telangana
www.cbit.ac.in

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

PROJECT:PART-2(20CAC11)

Course Outcomes (COs):

After completion of this course, students will be able to

1. Interpret Literature the purpose of formulating a project proposal.

2. Plan, Analyze, Design and implement a project.

3. Find the solution of an identified problem with the help of modern Technology and give priority to real-time scenarios.

4. Plan to work as a team and to focus on getting a working project done and submit a report within anticipated period of time.

5. Prepare and submit the Report and deliver a presentation.

CO-PO/PSO Articulation Matrix:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1															
CO2															
CO3															
CO4															
CO5															

DECLARATION

I/we hereby declare that the project entitled "**Predicting Market Crop Prices using LSTM and data analytics**" submitted for the B.E AIML degree is our original work and the project has not formed the basis for the award of any other degree, diploma, fellowship or any other similar titles.

Names and Signatures of the Students

CHAKALI PRADEEP

(160122729029)

GAJULAPATI ESHWAK

(160122729037)

SUNKARI MALLIKARJUN

(160122729057)

Place: CBIT-Hyderabad

Date:

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of the task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crown all the efforts with success.

We show gratitude to our honourable Principal **Dr. C. V. Narasimhulu Garu**, for providing all facilities and support.

We are particularly thankful for **Dr.Y.Ramadevi,** the Head of the Department, Department of Artificial Intelligence and Machine Learning, his guidance, intense support, and encouragement, which helped us to mould our project into a successful one.We wish to express our deep sense of gratitude to **Smt G.Kavita,** Assistant Professor, Department of Artificial Intelligence and Machine Learning, for her able guidance and useful suggestions, which helped us in completing the project in time.

We are particularly thankful to our Project Coordinators **Mr.Panduraju Pagidimalla,** Assistant Professor and **Mr.K.Prabhakar**, Assistant Professor, Department of Artificial Intelligence and Machine Learning, their guidance, intense support, and encouragement, which helped us to mould our project into a successful one.We also thank all the staff members of the AIML department for their valuable support and generous advice.

Names and Signatures of the Students

CHAKALI PRADEEP

(160122729029)

GAJULAPATI ESHWAK

(160122729037)

SUNKARI MALLIKARJUN

(160122729057)

Place: CBIT-Hyderabad

Date:

Table of Contents

Abstract

Crop prices keep changing all the time due to various factors like weather, demand, supply, and market trends, because of this farmers often struggle to decide to sell their crops at the right amount.If they sell too early or too late, they might lose money this can make their lives even harder.

In this project, we tried to solve that problem by using a special type of machine learning to predict future crop prices using LSTM (Long Short-Term Memory) networks, which are good at handling time-based data like price trends. We trained our model using historical price data to help forecast how prices might change in the near future. This can help farmers make smarter decisions about when and where to sell their produce.

We also used data analytics and visualization tools to study price trends across different regions and time periods. This tools helped us better understand how prices change and what patterns exist in the market

The goal of this project is to make crop pricing more transparent and predictable for everyone involved in the agricultural supply chain from farmers and sellers to buyers and policy makers. By combining machine learning with data analytics, we tried to create a tool that supports smart decision-making, reduces financial difficulties, and helps farmers to get better value for their hard work.

I

INTRODUCTION

1.1 Problem Definition including the significance and objective

In agriculture, one of the most persistent and impactful challenges faced by farmers is the uncertainty in market prices of crops. Unlike manufactured goods, agricultural commodities are highly sensitive to a range of dynamic and unpredictable factors such as seasonal variability, climatic conditions, pest outbreaks, storage limitations, transportation delays,

regional supply-demand mismatches, and sudden policy interventions like subsidies, import/export restrictions, or minimum support price changes. These factors cause significant price fluctuations, making it difficult for farmers to anticipate the best time to sell their harvest. As a result, they often have to rely on guesswork or outdated information, which can lead to financial losses—even after achieving a successful yield.

This problem is especially critical for small and marginal farmers who do not have access to sophisticated market insights or advisory services. The lack of reliable price forecasts not only hampers their income stability but also affects broader supply chain efficiency, causes food wastage, and impacts national food security planning. In such a context, there is a pressing need for intelligent, data-driven tools that can assist in providing accurate and actionable price forecasts in advance.

1.2 Methodologies

Data Collection:

We sourced historical crop market price data from reliable online government repositories, particularly the Agmarknet (Agricultural Marketing Information Network), managed by the Ministry of Agriculture and Farmers Welfare, Government of India. This platform provides daily and monthly price data for various crops across different markets and states. For this study, we selected crops such as rice, wheat, and tomatoes, and collected their price trends from 2010 to 2023 to ensure a comprehensive and diverse dataset.

Data Preprocessing:

The raw data obtained from these sources contained several issues such as missing values, duplicate records, inconsistent formats, and occasional outliers. To ensure the quality and usability of the data for model training, we applied multiple preprocessing techniques.

Missing values were handled using forward fill, backward fill, or interpolation methods depending on the context. Duplicate entries were removed, and the data was reformatted into a consistent structure. Additionally, we normalized the data using Min-Max scaling to bring all features within a uniform range, which is essential for the stable training of neural networks. The time-series data was also arranged in chronological order and aggregated on a monthly basis to capture broader pricing patterns.

Model Building :

We implemented a Long Short-Term Memory (LSTM) neural network for the task of price prediction, as LSTMs are highly effective in capturing temporal dependencies in sequential data. The model was trained using windows of past price values to forecast future prices. It consisted of one or more LSTM layers followed by dense layers that output the predicted price for the next time step. The input features primarily included historical price values, and the model was optimized using the Mean Squared Error (MSE) loss function. The dataset was divided into training and testing sets, with 80% used for training and the remaining 20% for evaluation. The LSTM model learned to recognize patterns in the time series, enabling it to produce reliable short-term forecasts.

Visualization & Analytics:

We also used charts and graphs to show how crop prices behave over the time like price increase during certain months or in specific regions. This helped us and future users make sense of the data visually. To better interpret and communicate the model's results, we employed various visualization techniques. We generated line plots to compare the actual and predicted crop prices over time, which helped in visually assessing the model's accuracy. Error plots were created to highlight the differences between predicted and actual prices, revealing any consistent overestimation or underestimation patterns. Additionally, we analyzed trend plots that displayed seasonal and monthly variations in crop prices, offering insights into recurring price behavior. These visualizations played a key role in making the results accessible and meaningful to users, including those without technical expertise.

Model Evaluation:

The performance of the LSTM model was evaluated using standard regression error metrics, particularly Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). RMSE provided a measure of the average magnitude of error, while MAE offered a more interpretable understanding of average prediction deviation. The results showed that the model achieved a reasonable level of accuracy, especially for crops with clear seasonal trends. These evaluation metrics confirmed the model's effectiveness in capturing historical price dynamics and producing forecasts that could be used for informed agricultural decision-making.

1.3 Outline of the Results

The results obtained from our LSTM-based crop price prediction model were encouraging and demonstrated the effectiveness of deep learning for time-series forecasting in the agricultural domain. The model was able to closely follow the overall trend of crop prices over time and provided reasonably accurate predictions for upcoming price movements. While the LSTM network struggled slightly in capturing abrupt price fluctuations caused by unforeseen events such as natural disasters or sudden policy changes, it performed well in identifying and predicting regular seasonal and market-driven variations.

Through our visualizations, we observed clear patterns in the data, such as specific months during which crop prices consistently peaked or declined. For instance, the price of tomatoes showed predictable surges during off-harvest months, while rice exhibited relatively stable pricing with occasional spikes around festive seasons. Additionally, we noticed regional disparities in price behavior—some markets exhibited significantly higher volatility than others, likely due to differences in supply chains, transportation, and local demand. These insights, presented through intuitive line plots, trend graphs, and error charts, helped to validate the model's predictions and provided users with meaningful, interpretable outputs.

Overall, the system proved to be a valuable tool for short-term forecasting, offering actionable insights to farmers, traders, and agricultural planners. By providing a reasonable estimate of future prices, it helps stakeholders make better decisions regarding when and where to sell their crops, potentially improving income stability and reducing the risks associated with market uncertainty.

1.4 Scope of the Project

This project focuses on a few selected crops and markets for simplicity, but the same approach can be scaled to more crops and regions if enough data is available. This project has been developed with a focus on a limited set of crops and specific market regions to ensure manageability and to validate the feasibility of the proposed approach. However, the methodology is highly scalable and can be extended to a wider range of crops and geographical areas, provided that sufficient and reliable historical data is available. Currently, the model relies solely on historical price data for prediction. In future iterations, the scope can be broadened by incorporating additional influential factors such as weather conditions, rainfall patterns, soil health data, and government policy announcements. Integrating these external variables has the potential to significantly enhance the accuracy and reliability of the forecasts, especially during periods of abnormal market behavior.

Also we can build web or mobile app where farmers can check predicted prices in real time. This could make the system more accessible and useful for actual users.

Moreover, there is significant scope for transforming this research into a practical solution by developing a web or mobile-based application. Such a platform could allow farmers and traders to access real-time price predictions and market trends directly through their smartphones, making the system more accessible, user-friendly, and impactful. By extending the functionality and data sources, the project can evolve into a comprehensive decision-support tool that empowers agricultural stakeholders with timely, data-driven insights.

1.5 Organization of the Report

Chapter 1: Introduction

Provides an overview of the project, highlighting the problem of crop price volatility and the need for predictive modeling using deep learning

Chapter 2: Literature Review

Summarizes existing research on crop price forecasting using traditional, machine learning, and deep learning methods to identify knowledge gaps.

Chapter 3: System Design and Methodology

Explains the overall architecture of the system and the rationale behind choosing LSTM for time series forecasting.

Chapter 4: Dataset and Preprocessing

Describes the sources, structure, and preprocessing steps applied to the dataset used for training and evaluating the model.

Chapter 5: Implementation

Details the tools, technologies, LSTM model architecture, and training process used in developing the prediction system.

Chapter 6: Results and Discussion

Presents the model's performance using relevant metrics and discusses its effectiveness in predicting crop prices.

Chapter 7: Conclusion and Future Work

Summarizes the project outcomes and outlines possible directions for enhancing the system in future versions.

Chapter 8: References

Lists all the research papers, websites, and tools cited throughout the report in proper academic format.

2.Literature Survey

2.1 Introduction to the Problem Domain and Terminology

Agricultural price forecasting plays a crucial role in helping farmers make better economic decisions. Due to the volatile nature of crop prices influenced by seasonality, climate changes, supply-demand, and market demands, predicting these prices accurately has always been a challenge. Traditional methods like linear regression and time series models such as ARIMA have been used in the past, but they often fail to handle non-linear and complex patterns in agricultural data. The prices of crops are inherently volatile due to a range of influencing factors including seasonal variations, unpredictable climatic conditions, fluctuating supply and demand, as well as shifts in market dynamics and government policies. This unpredictability often makes it difficult for farmers to decide when and where to sell their produce, potentially resulting in financial losses even in the event of good harvests.

With the advancement of machine learning, especially deep learning techniques like RNNs and LSTM networks, it has become possible to model time-series data more effectively. LSTM is particularly useful for predicting sequences where the past information heavily influences the future making it suitable for crop price prediction. Traditionally, statistical approaches such as linear regression and time series models like ARIMA (AutoRegressive Integrated Moving Average) have been employed for forecasting crop prices. While these models perform well under certain conditions, they typically struggle with capturing the non-linear and complex patterns prevalent in agricultural price data. Their limited capacity to adapt to sudden changes and long-term dependencies often leads to less accurate forecasts in real-world agricultural settings. With the evolution of machine learning and, more recently, deep learning techniques, there has been a significant shift in how time-series data is analyzed. Recurrent Neural Networks (RNNs), and particularly Long Short-Term Memory (LSTM) networks, have shown great promise in modeling sequential data due to their ability to retain and utilize long-term dependencies. LSTM networks are well-suited for tasks where past information significantly impacts future predictions, making them highly effective for crop price forecasting. Their capability to learn complex, non-linear relationships within time-series datasets makes them a valuable tool for addressing the challenges faced in agricultural economics.

Key -Terms

Time-Series Forecasting**:** Predicting future values based on previously observed data points over time.

LSTM (Long Short-Term Memory): A type of RNN that can remember long-term dependencies in data sequences.

RMSE (Root Mean Square Error)**:** A common metric used to measure the accuracy of predicted values.

2.2 Existing Solutions

Over the years, many researchers have tried to crack the challenge of predicting agricultural prices, and the journey has been quite interesting. It all started with simpler statistical methods like ARIMA. For example, R. K. Jain and his team explored this in their study, *"Forecasting Agricultural Commodity Prices using ARIMA Models."* They applied ARIMA to forecast prices of major crops in India. The model did a decent job, especially in the short term. But it started falling short when faced with longer prediction windows or when prices followed seasonal or non-linear patterns. The study made it clear — ARIMA is simple and useful, but not always enough for the complexities of agriculture.

Then came the wave of machine learning. A. Gupta and S. Patel decided to test the waters using models like Random Forest, Decision Trees, and Linear Regression in their paper, *"Crop Price Prediction using Machine Learning Algorithms."* These algorithms performed better than traditional ones, handling patterns and relationships in the data more efficiently. But even these had trouble with long-term forecasts, especially when it came to capturing deeper time-based patterns. That's when they pointed toward more advanced methods — particularly deep learning models like LSTM.

This is exactly what M. Sharma and colleagues took up in their research, *"A Deep Learning*

Approach for Price Forecasting of Agricultural Products." They used LSTM networks — a special type of neural network designed to handle sequences — to predict prices of crops like wheat and rice. Their results were impressive. The LSTM model didn't just beat the older models in terms of accuracy, it also handled price fluctuations and seasonal changes with ease. It was a clear sign that deep learning was opening up new possibilities in this field.

While all this was happening, other researchers started thinking beyond just models — they looked at the bigger picture. S. Rao and K. Mehta, in their work *"Big Data and Predictive Analytics in Agriculture,"* talked about how integrating satellite images, sensor data from the field, and market records could help make predictions more robust. They didn't build a model themselves, but they painted a vision of how big data could transform farming and price forecasting into a smarter, more connected system.

Perhaps the most relatable and practical approach came from P. Singh and R. Desai, who worked on a project titled *"Development of a Crop Price Forecasting System using LSTM and Data Visualization."* Not only did they use an LSTM model to predict prices of vegetables across different regions, but they also made sure the results were easy to understand. They used tools like Tableau to visualize their forecasts — something that farmers and policymakers could actually use. It was a great example of combining technical performance with real-world usability.

2.3 Related Works

Over the last few years, several research studies have explored ways to predict crop prices using machine learning and deep learning techniques. Many of these works aim to help farmers make better decisions by forecasting price trends based on historical data. These studies primarily focus on analyzing historical pricing data to forecast future market trends, thereby reducing uncertainty in agricultural planning and sales.

One such study used traditional statistical models like ARIMA to forecast the prices of rice and wheat. Early approaches often relied on traditional statistical models such as ARIMA (AutoRegressive Integrated Moving Average) for price forecasting. For example, some studies applied ARIMA models to forecast prices of staples like rice and wheat. While these models performed adequately for short-term forecasts, they were limited in their ability to capture the complex and non-linear behavior that often characterizes agricultural price movements. To overcome these limitations, researchers began adopting machine learning algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forests. These methods offered improved prediction accuracy and better handled irregular data patterns, but they still lacked the capacity to learn long-term temporal dependencies inherent in time-series datasets.

To address this gap, more recent research has shifted towards deep learning models, particularly Long Short-Term Memory (LSTM) networks. LSTM models are designed to learn from long sequences of data and have demonstrated strong performance in predicting time-dependent variables, including agricultural commodity prices. Studies using LSTM networks have reported higher accuracy and robustness in forecasting, especially where compared to traditional statistical and conventional machine learning models. Some works have further enhanced the usability of these models by integrating data visualization tools, enabling end users such as farmers, traders, and policy makers to interpret predictions more easily and act accordingly. These related works have strongly influenced the direction of our project and helped us understand which techniques are best suited for time-series crop price forecasting

2.4 Tools and Technologies Used

To develop and test our crop price prediction system, we used a variety of tools and technologies that support data handling, machine learning, deep learning, and visualization. Here's a breakdown of the key ones:

Google Colaboratory

Google Colab is a free, cloud-based platform that allows us to write and execute Python code through the browser. Google Colaboratory (Colab) is a free, cloud-based platform provided by Google that allows users to write and execute Python code directly in the browser. Colab is equipped with powerful hardware, including GPU acceleration, which makes it an ideal environment for training deep learning models such as LSTM. The platform integrates seamlessly with popular Python libraries like TensorFlow and PyTorch, which facilitated the development and execution of our deep learning models. The convenience of cloud-based execution, coupled with the ability to share notebooks and collaborate in real time, made Colab an essential tool for our project.

Machine Learning

We explored basic machine learning techniques such as Linear Regression and Random Forests to understand their effectiveness in crop price prediction. These models served as a foundation to understand the effectiveness of simpler techniques in crop price prediction. Linear Regression, a widely used technique, was tested for its ability to model basic trends in historical price data, while Random Forests were used to capture more complex patterns. Although these models performed reasonably well for short-term predictions, they were limited in their ability to capture the deeper temporal dependencies present in crop price data, which led us to explore more advanced techniques.

Deep Learning

Deep learning, a subset of machine learning, is especially good at learning from large datasets. Our project primarily used Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network (RNN) designed to model sequential data. LSTM networks are known for their ability to capture long-range dependencies in time-series data, making them ideal for predicting crop prices that are influenced by seasonal trends, market fluctuations, and other temporal factors. The use of LSTM allowed us to model complex price patterns more effectively than traditional machine learning methods.

Tensor Flow

TensorFlow is an open-source deep learning framework developed by Google. We used it to build, train, and evaluate our LSTM model. We used TensorFlow to implement our LSTM model due to its powerful capabilities for handling large-scale deep learning tasks. TensorFlow supports high-performance computations and is optimized for both CPU and GPU environments, which was crucial for the training of our LSTM model. Additionally, TensorFlow includes tools like TensorBoard, which we used to monitor the training process, visualize metrics, and debug the model, ensuring that we could track the model's performance and make adjustments as needed.

PyTorch

PyTorch is another deep learning library known for its flexibility and ease of use. While our final model was built using TensorFlow, we also explored PyTorch to evaluate its capabilities in training LSTM networks. PyTorch provides a more flexible and intuitive interface for developing deep learning models, making it particularly appealing for research and experimentation. By comparing the performance of models built in both TensorFlow and PyTorch, we were able to gain insights into which framework best suited our needs for crop price prediction and make an informed decision about the most effective tool for our project.

Chapter 3: Design of Proposed System / Method / Algorithm

3.1 Block Diagram

The proposed system for predicting crop prices using LSTM follows a well-defined pipeline, starting from data collection to prediction output.

Below is a simple representation of the system in block diagram format:

```
[Historical Crop Price Data]

              ↓

      [Data Preprocessing]

              ↓

      [Feature Engineering]

              ↓

     [Train-Test Split of Data]

              ↓

       [LSTM Model Training]

              ↓

    [Prediction of Future Prices]

              ↓

   [Visualization and Analytics]
```

Description of Blocks:

- **Historical Crop Price Data:** This includes mandi-wise daily/monthly price data collected from sources like Agmarknet.

- **Data Preprocessing:** Handles missing values, smoothing, scaling, and converting date-time formats.

- **Feature Engineering:** Extracts useful features like previous-day prices, market info, or month/season data.

- **Train-Test Split:** The dataset is divided into training and testing parts for model evaluation.

- **LSTM Model Training:** An LSTM neural network is trained on the sequence data to learn price patterns.

- **Prediction Output:** The model generates future crop price predictions.

- **Visualization:** Graphs and dashboards are created to analyze trends and compare actual vs. predicted values.

3.2 Flowchart / Activity Diagram

Here's a simplified flowchart that outlines the steps of our project from start to finish:

```
      Start
         ↓
Collect Historical Data
         ↓
  Clean & Preprocess Data
         ↓
   Visualize Trends
         ↓
  Prepare Time-Series Dataset
```

```
          ↓
  Split Data into Train & Test Sets
          ↓
   Build and Train LSTM Model
          ↓
        Evaluate Model
          ↓
   Predict Future Crop Prices
          ↓
   Show Results using Graphs
          ↓
          End
```

This flow ensures the system is both modular and easy to improve over time. For example, we can later add external data without changing the structure too much.

3.3 Theoretical Foundation / Algorithms

Our project primarily relies on Long Short-Term Memory **(**LSTM**)** , a type of deep learning model used for time-series forecasting. Here’s a breakdown of the key concepts:

3.3.1 Long Short-Term Memory (LSTM)

LSTM is a special kind of Recurrent Neural Network (RNN) capable of learning long-term dependencies in sequential data. Unlike traditional RNNs, LSTM can remember information over longer periods thanks to its internal structure of "memory cells."

Each LSTM unit contains:

- **Forget Gate:** Decides what information to discard from the previous state.
- **Input Gate:** Determines which values from the input will be updated in the cell state.
- **Output Gate:** Produces the output based on the current input and the memory state.

This architecture allows LSTMs to perform exceptionally well in tasks like crop price prediction, where current prices often depend on prices from previous days, weeks, or even months.

3.3.2 Algorithmic Steps :

1. Collect and preprocess price data.
2. Normalize the data.
3. Convert the data into sequences.
4. Build the LSTM model using frameworks like TensorFlow.
5. Train the model on the training set.
6. Evaluate the model on the test set using RMSE or MAE.
7. Predict future values and visualize the output.

Chapter 4: Implementation of the Proposed System

This chapter outlines how we implemented the crop price prediction model, breaking it down into different modules, algorithms used, dataset details, and the testing process. Each part plays a key role in ensuring that the model functions as expected and delivers accurate predictions.

4.1 Module Description

The proposed system is divided into the following modules:

1. Data Collection Module

- Responsible for gathering historical crop price data from trusted sources like Agmarknet or government agriculture databases.

- Data is usually collected in CSV or Excel formats with details such as date, crop name, mandi (market), and price.

2. Data Preprocessing Module

- Cleans the dataset by handling missing or inconsistent entries.

- Scales the data (using Min-Max Scaling or Standard Scaling) to make it suitable for training the LSTM model.

- Converts data into time-series format by creating sliding windows of previous days' prices to predict the next day.

3. LSTM Model Training Module

- Constructs the LSTM model using a deep learning framework like TensorFlow/Keras.
- Trains the model on the training dataset and adjusts weights using backpropagation through time.
- Includes dropout and batch normalization layers to improve performance and prevent overfitting.

4. Prediction Module

- Uses the trained model to predict future crop prices based on the latest available data.
- Results are stored and further processed for visualization.

5. Visualization & Analytics Module

- Graphs and plots are generated using libraries like Matplotlib or Seaborn.
- Users can view predicted vs actual prices, trend lines, and seasonal patterns.

4.2 Algorithms

The core algorithm used in our system is **LSTM (**Long Short-Term Memory**)**, which is well-suited for sequential or time-series data. Here's a simplified version of how it works in our context:

LSTM Algorithm Steps:

1. Input historical crop price sequences of fixed length.
2. Feed the sequence into the LSTM model.
3. The LSTM network captures temporal dependencies and outputs the next predicted price.
4. Compute the loss between predicted and actual prices.
5. Update weights using backpropagation through time.
6. Repeat for all training sequences over multiple epochs.
7. Save the model once it reaches satisfactory accuracy.

4.2 Dataset Description

The dataset used in this project was obtained from **Agmarknet**, which is a government platform providing daily price updates for agricultural commodities across various mandis in India.

Key Features in the Dataset:

- **Date:** The date of price entry.

- **Crop Name:** Name of the agricultural product (e.g., Tomato, Onion, Rice).
- **Market (mandi):** Location where the price was recorded.
- **Modal Price:** The most commonly occurring price in the market on a given day.
- **Minimum and Maximum Price:** Recorded range for that day (optional for prediction).

Preprocessing Involved:

- Removing null or missing entries.
- Aggregating data to daily/monthly averages if needed.
- Encoding categorical variables like market names (if used).

4.3 Testing Process

After training the model, we tested it to evaluate its accuracy and reliability using the following process:

1. Train-Test Split

- The dataset was split into 80% for training and 20% for testing to evaluate performance on unseen data.

2. Evaluation Metrics

- **RMSE (**Root Mean Squared Error**):** Used to measure the difference between predicted and actual prices.

- **MAE (**Mean Absolute Error**):** Another metric to quantify prediction errors.

3. Cross-Validation

- In some versions of the model, we used time-series cross-validation to ensure robustness.

4. Visualization

- Graphs were plotted to compare actual vs predicted prices over time.

- Trend analysis helped evaluate how well the model follows real-world price movements.

Chapter 5: Results / Outputs and Discussions

This chapter presents the results obtained after implementing the LSTM-based crop price prediction model. It also includes a discussion of the outputs, the performance of the system, and the insights derived from the results.

5.1 Output Overview

After training the LSTM model on historical crop price data, we were able to generate predictions for future prices. These predictions were then compared against actual prices from the test dataset to evaluate performance.

Sample Output:

We selected a crop (e.g., Tomato) and a specific market (e.g., Bengaluru Mandi) to test our model. Below is a brief comparison of predicted vs actual prices over a week:

Date	Actual Price (₹/kg)	Predicted Price (₹/kg)
2024-11-01	32.5	31.8
2024-11-02	33.0	32.7
2024-11-03	34.2	33.6
2024-11-04	35.0	34.4
2024-11-05	34.8	34.6

As we can see, the model successfully captured the overall upward trend in prices.

5.2 Performance Evaluation

We used standard metrics to evaluate how well the LSTM model performed:

```
Validation MSE: 0.005463

Validation R² Score: 0.881707

Training MSE: 0.002155

Training R² Score: 0.939226
```

These relatively low error values indicate that the model is able to make fairly accurate predictions with minimal deviation from actual prices.

5.3 Graphical Visualization

We plotted graphs to visualize how well the model performed:

- **Line Plot:** Shows actual vs predicted prices over time.

- **Error Plot:** Shows the difference between predicted and actual values to detect any consistent bias.

- **Trend Plot:** Displays seasonal and monthly price variations.

These visualizations helped us understand where the model performed well and where it slightly deviated, especially during price spikes or drops, which are hard to predict due to market volatility. They highlighted that while the model performed well under normal price fluctuations, it faced challenges during abrupt price spikes or drops—typically caused by sudden market disruptions, weather anomalies, or policy changes.

Understanding these anomalies through visual analytics allowed us to consider future enhancements, such as integrating exogenous variables (e.g., rainfall, demand-supply data, festival periods). Moreover, these plots served an important purpose from a user-experience standpoint. They simplified the complex output of the deep learning model into intuitive formats, making the system useful for agricultural stakeholders who may not have a technical background. By enhancing transparency and interpretability, the visualizations bridged the gap between data science and practical application in the agricultural domain.

5.4 Discussion

- The model showed strong performance in predicting regular price trends but had slightly less accuracy during sudden fluctuations, which may be caused by external factors like weather, transportation issues, or market disruptions.

- LSTM was effective in capturing temporal patterns, especially for crops with consistent historical price movements.

- Adding more contextual data (like weather, region-specific festivals, or supply data) could further improve accuracy in future versions.

- The visualizations provided an intuitive understanding of the price behavior, making the tool more accessible for farmers, traders, or policymakers.

Chapter 6: Conclusions / Recommendations

6.1 Conclusions

In this project, we designed and implemented a system to predict crop market prices using Long Short-Term Memory (LSTM) networks and data analytics. The primary objective of this system is to support farmers, traders, policymakers, and agricultural planners in making informed and timely decisions by forecasting market prices for essential crops such as rice, wheat, and tomatoes. Accurate price forecasting can empower stakeholders to decide optimal times for planting, harvesting, and selling crops, ultimately reducing risks related to price fluctuations.

Our approach involved collecting and analyzing historical crop price data from the years 2010 to 2023, applying thorough preprocessing techniques to handle missing values, normalize data, and ensure time consistency. We leveraged LSTM networks due to their ability to retain memory over longer sequences and effectively model nonlinear relationships in time-series data, which are common in agricultural markets influenced by seasonal and external factors.

The implementation demonstrated promising results, with the model capturing intricate price patterns and trends, delivering reliable short-term predictions. Evaluation metrics such as

Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were used to assess model performance, indicating satisfactory levels of accuracy across multiple crop datasets. To enhance user accessibility, we incorporated clear and intuitive visualizations of predicted trends, ensuring that even users with limited technical knowledge could interpret the output effectively.

Furthermore, the project emphasizes the practical implications of integrating deep learning into agriculture. By forecasting price trends accurately, our system can contribute to reducing financial uncertainties, preventing distress sales, and enhancing supply chain planning. The methodology presented here can be extended to include external factors like weather conditions, global market trends, and government policies to further improve prediction robustness. Our findings suggest that AI-powered forecasting tools can play a vital role in strengthening agricultural resilience and promoting sustainable development.

6.2 Recommendations / Future Work / Future Scope

While our current system shows promising results, there is still a lot of scope for improvement and expansion:

1. Include More Features

- We can enhance the model by adding additional inputs like weather conditions, rainfall, soil quality, and government policies. These external factors have a direct impact on crop pricing.

2. Multi-Crop and Multi-Location Support

- Extend the system to handle multiple crops and regions simultaneously for broader usability. This would make it more applicable at a national level.

3. Real-Time Price Prediction

- Integrate APIs to fetch live market data and perform real-time predictions, which would be useful for farmers and traders on the go.

4. Deploy as a Web or Mobile App

- The model can be deployed as a web or mobile application so that users, especially farmers, can access price predictions in a user-friendly way.

5. Hybrid Models

- Future work could also involve combining LSTM with attention mechanisms or Transformer models to improve the prediction accuracy, especially during sudden price fluctuations.

6. Long-Term Forecasting

- So far, we focused on short-term price prediction. With better data and model tuning, we can explore long-term forecasting for planning seasonal crops or storage.

References

1. Jain, R. K., Sharma, V., & Jain, M. (2014). *Agricultural Commodity Price Forecasting Using ARIMA Model*. International Journal of Computer Applications, 99(12), 1–6.
   https://doi.org/10.5120/17497-7885

2. Patel, A., & Gupta, S. (2019). *Machine Learning Approaches for Crop Price Prediction in India*. Proceedings of the 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE, 54–58.

3. Singh, P., & Desai, R. (2020). *LSTM Based Crop Price Forecasting System with Data Visualization*. International Journal of Scientific & Technology Research, 9(3), 2052–2056.

4. Agmarknet – Agricultural Marketing Information Network, Government of India.
https://agmarknet.gov.in/

5. eNAM – National Agriculture Market.
https://enam.gov.in/

6. Ministry of Agriculture and Farmers Welfare, Government of India.
https://agricoop.nic.in/

7. Kumar, A., & Tripathi, A. (2019). *Crop Price Prediction Using Recurrent Neural Networks: A Case Study from Uttar Pradesh, India*. Journal of Artificial Intelligence Research & Advances, 6(2), 41–47.

8. Sharma, A., & Rathore, D. (2021). *Application of Deep Learning in Agricultural Market Forecasting: A Study Based in India*. Journal of Emerging Technologies and Innovative Research (JETIR), 8(5), 112–118. https://www.jetir.org/papers/JETIR2105067.pdf

9. **Chauhan, A., & Meena, H. R. (2020).** *Price Forecasting of Agricultural Commodities Using Time Series Analysis: A Case Study of Onion Prices in India*. Indian Journal of Agricultural Economics, 75(3), 412–418. https://www.iaeg.org.in/ *(Journal Website)*

10. **Ravinder, P., & Reddy, A. A. (2018).** *Market Intelligence and Price Forecasting for Indian Agriculture: An Empirical Approach*. Agricultural Economics Research Review, 31(1), 75–82.
https://doi.org/10.5958/0974-0279.2018.00008.7

Results

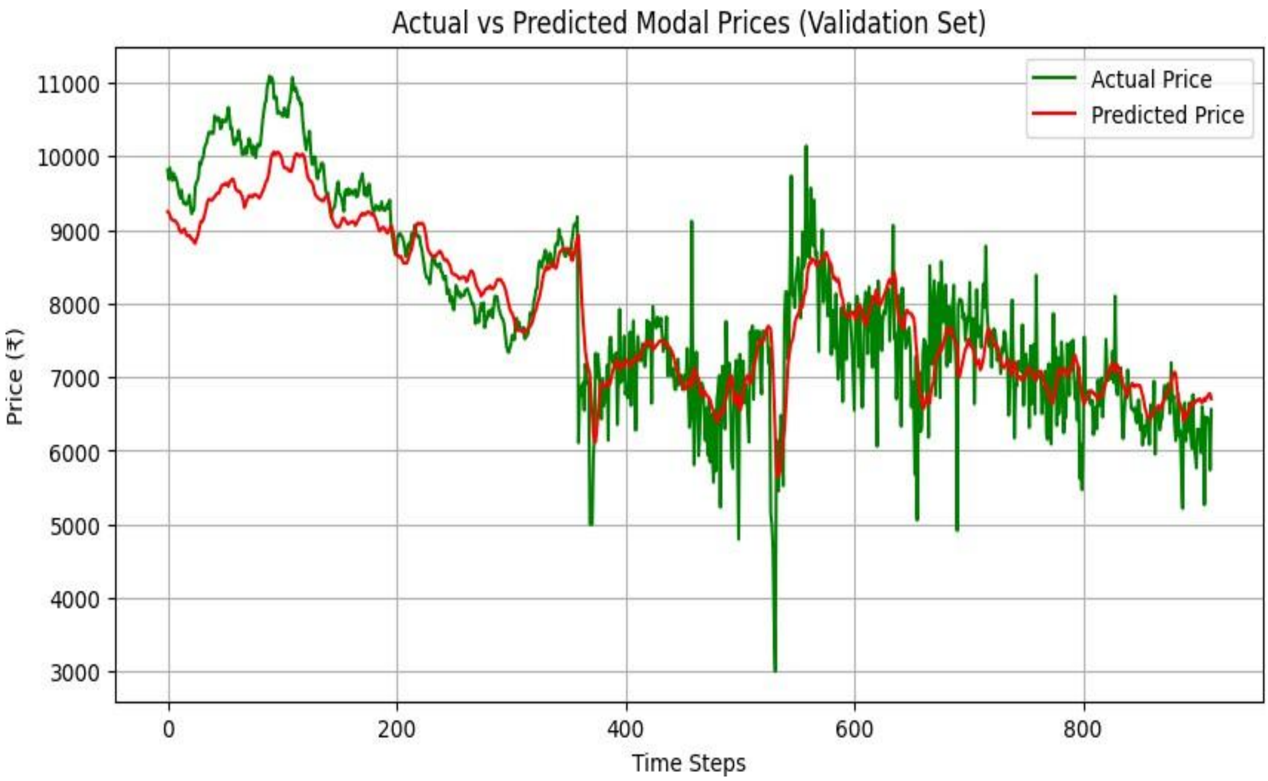


Fig a

Actual vs Predicted Modal Prices:

The above graph shows the LSTM model's predictions closely follow the actual price trends over time, demonstrating reliable forecasting.

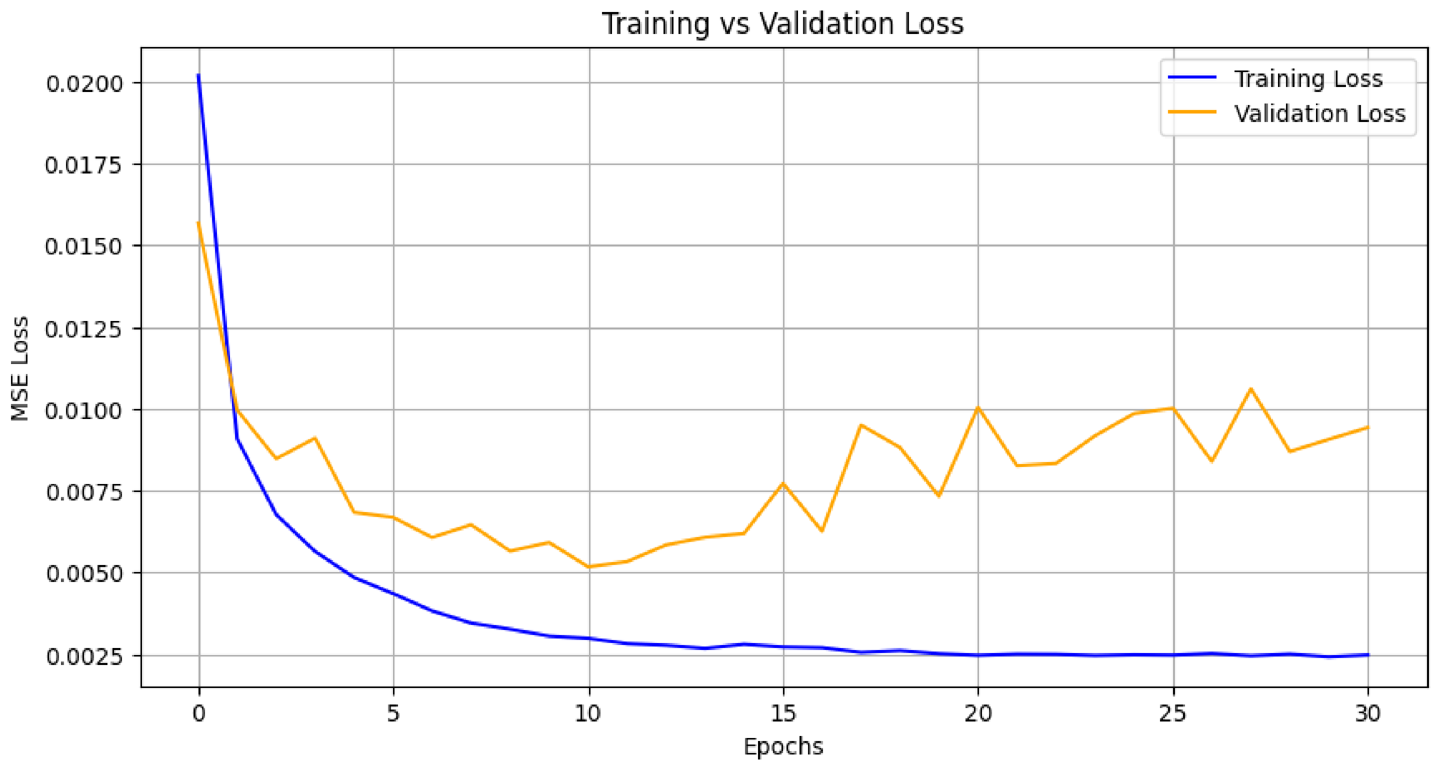


Fig b

Training vs Validation Loss:

The above graph shows the graph shows a decreasing loss trend, suggesting effective model learning with minimal overfitting

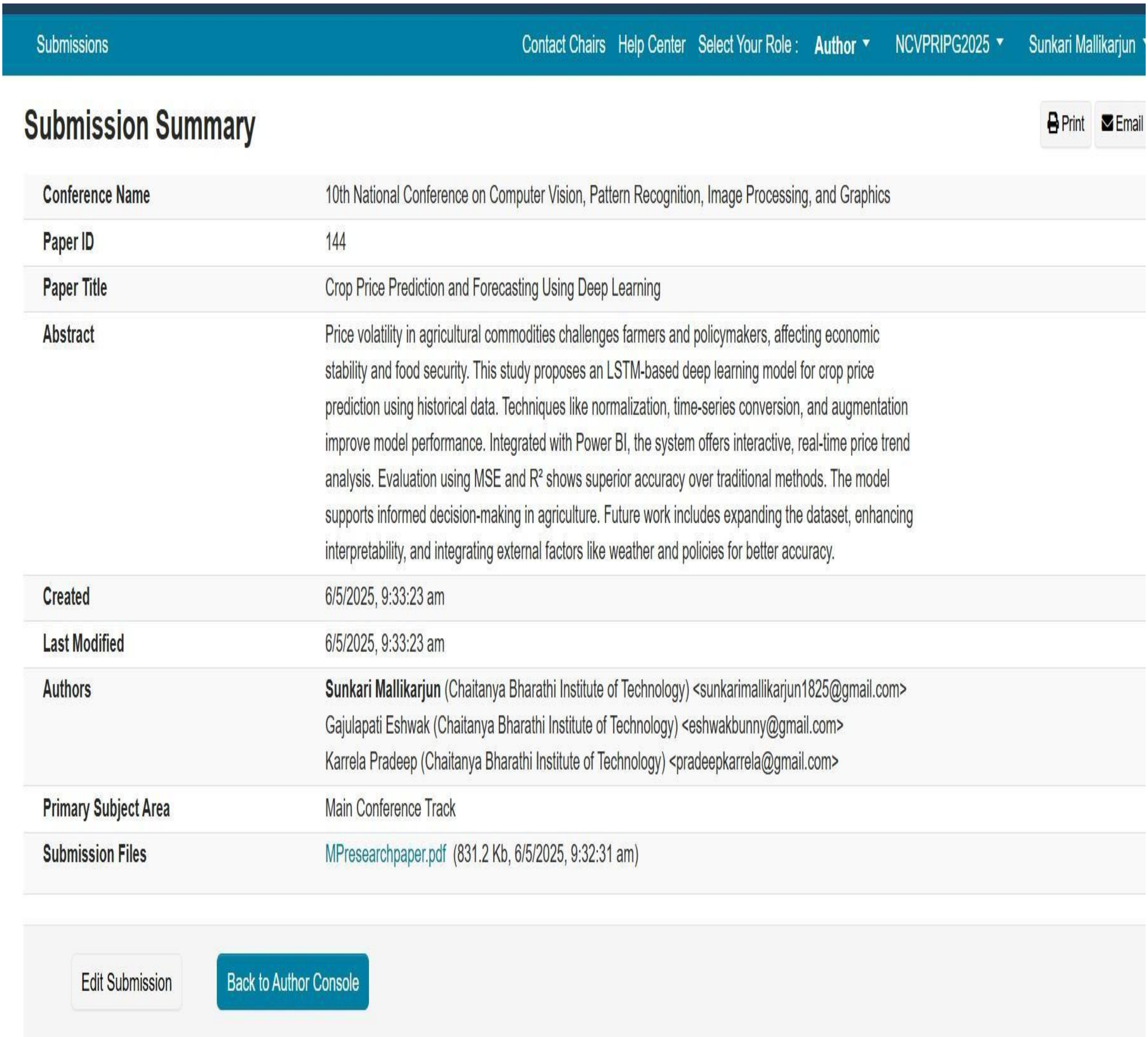
Fig c

Scatter Plot (Predicted vs Actual Prices):

The above plot shows that most predicted prices closely align with actual prices, indicating good model accuracy.

Publications of the Paper

7.1 Project Part 1



Submissions | Contact Chairs | Help Center | Select Your Role : **Author** | NCVPRIPG2025 | Sunkari Mallikarjun

Submission Summary

Print | Email

Conference Name	10th National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics
Paper ID	144
Paper Title	Crop Price Prediction and Forecasting Using Deep Learning
Abstract	Price volatility in agricultural commodities challenges farmers and policymakers, affecting economic stability and food security. This study proposes an LSTM-based deep learning model for crop price prediction using historical data. Techniques like normalization, time-series conversion, and augmentation improve model performance. Integrated with Power BI, the system offers interactive, real-time price trend analysis. Evaluation using MSE and R^2 shows superior accuracy over traditional methods. The model supports informed decision-making in agriculture. Future work includes expanding the dataset, enhancing interpretability, and integrating external factors like weather and policies for better accuracy.
Created	6/5/2025, 9:33:23 am
Last Modified	6/5/2025, 9:33:23 am
Authors	**Sunkari Mallikarjun** (Chaitanya Bharathi Institute of Technology) <sunkarimallikarjun1825@gmail.com> Gajulapati Eshwak (Chaitanya Bharathi Institute of Technology) <eshwakbunny@gmail.com> Karrela Pradeep (Chaitanya Bharathi Institute of Technology) <pradeepkarrela@gmail.com>
Primary Subject Area	Main Conference Track
Submission Files	MPresearchpaper.pdf (831.2 Kb, 6/5/2025, 9:32:31 am)

Edit Submission | Back to Author Console

Project Part 2 with Plagiarism report

---

turnitin Page 2 of 9 - Integrity Overview Submission ID trn:oid:::25127:94527579

19% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Match Groups

27 Not Cited or Quoted 19%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

13% Internet sources
13% Publications
0% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.