## ACKNOWLEDGEMENT

I would like to extend my heartfelt gratitude to DLithe Consultancy Services Private Limited for providing me with the invaluable opportunity to undertake my internship at their esteemed institution. Their support and guidance were instrumental in shaping my project, and I am truly grateful for the experience.

I would also like to express my deepest thanks to Ms. Medini, who served as an exceptional guide throughout this internship journey. Her expertise, encouragement, and mentorship were pivotal in making this experience rewarding and enriching.

Additionally, I am grateful to Mangalore Institute of Technology & Engineering for giving me the chance to pursue this internship at DLithe Consultancy Services Private Limited. The college's support and encouragement have been crucial in enabling me to gain practical industry experience and further my learning.

Finally, I want to acknowledge all the individuals who assisted me in various ways, providing valuable insights and contributing to my growth and knowledge during the internship period. Thank you all for your unwavering support.

## ORGANIZATIONAL INFORMATION

DLithe is a technology-based product company that has been serving IT companies and academic institutions since the year 2018. The company is led by industry professionals with two decades of experience. For IT companies, DLithe offers services such as technology consultancy, project development, IT recruitment, staffing, competency development, and content development. On the other hand, the company serves academic institutions by providing competency development services in niche technologies like artificial intelligence, internet of things, robotics, cybersecurity, augmented reality, and more. DLithe has also developed the arm-based Cortex M3 series microcontroller and the ioCube product in the embedded and IoT domain.

During my rewarding internship in the field of Artificial Intelligence and Machine Learning, I had the privilege of being a part of an exceptional program under the guidance of this renowned organization. Throughout the internship, I gained comprehensive insights into diverse industry verticals, spanning from understanding project requirements to the final deployment phase.

DLithe’s internship program provided me with a valuable opportunity to immerse myself in real-world scenarios, gaining exposure to industry best practices and learning how to implement AI and ML solutions within an agile project life cycle. The supportive environment and dedicated mentors at the organization ensured that I could explore practical use cases for AI and ML implementation, enabling me to grow and learn during insightful post-mentoring sessions.

One notable aspect of the internship was the opportunity to work on real-world project, including a crop price prediction project. DLithe provided guidance and mentoring throughout the project, allowing me to gain hands-on experience with different machine learning models and their application in solving practical problems like crop price prediction.

This internship experience has equipped me with a strong foundation in artificial intelligence and machine learning, positioning me well for a career in this dynamic and rapidly evolving field. Overall, this internship has been a transformative experience, equipping me with not only technical skills but also a deeper understanding of how AI and ML technologies play a vital role across various industries.

**CONTENTS**

### ACKNOWLEDGEMENT 2

### ORGANIZATIONAL INFORMATION 3

### CONTENTS 4

### ABSTRACT 5

### INTERNSHIP OBJECTIVES 6

### WEEKLY OVERVIEW OF INTERNSHIP ACTIVITIES 8

### CHALLENGES & LEARNING OUTCOMES 10

### PROJECT DETAILS

### INTRODUCTION 12

### LITERATURE SURVEY 13

### PROBLEM STATEMENT 15

### PROJECT OBJECTIVES 16

### METHODOLOGIES 18

### IMPLEMENTATION 20

### RESULT AND FUTURE SCOPE 22

### APPENDIX 25

### BIBLIOGRAPHY 27

## ABSTRACT

This report presents a comprehensive study on the development and application of Long Short-Term Memory (LSTM) neural networks for the prediction of modal prices of two critical agricultural commodities, Arecanut (Coca) and Coconut (Grade-I). Modal prices represent a crucial economic indicator in the agricultural market, signifying the price range within which different merchants transact their produce. Understanding and accurately predicting modal prices are of utmost importance for both farmers and merchants to make informed decisions about their trading strategies.

In this study, a dataset spanning the last seven years was collected and preprocessed to train and validate the LSTM model. The data consisted of daily price records for Arecanut and Coconut, including minimum and maximum prices. The LSTM model was designed to take into account sequences of prices over a three-day window, allowing it to capture temporal dependencies and patterns in the data.

Crop price prediction plays a crucial role in agricultural planning and decision-making. In this study, we propose a neural network-based approach to forecast crop prices. The neural network model utilizes historical price data and other relevant factors as input features to predict future crop prices. We preprocess and clean the data, perform feature engineering, and train the neural network using a deep learning framework. Our results demonstrate the effectiveness of the neural network in capturing complex price patterns and trends.

we analyze the importance of different input features and investigate the model's ability to adapt to changing market conditions. Our findings highlight the neural network's adaptability and its capability to provide valuable insights for farmers, traders, and policymakers in the agricultural sector. Overall, this research contributes to the field of agricultural economics by introducing a robust and data-driven approach to crop price prediction, enhancing decision-making processes and improving resource allocation in the agricultural industry.

## INTERNSHIP OBJECTIVES

The primary objectives of the AI and ML internship was designed to equip us with a comprehensive skill set and practical knowledge in various areas of Artificial Intelligence and Machine Learning. The key objectives included:

* **Learning Python Basics:** The internship aimed to provide a strong foundation in Python programming, as it is one of the most widely used languages in AI and ML. Participants were introduced to Python syntax, data structures, and essential libraries used in AI and ML development.
* **Gain Practical Experience:** The primary goal of this internship was to gain practical, hands- on experience in the field of artificial intelligence and machine learning. This involved working on real-world projects and applying AI and ML concepts to solve practical problems.
* **Understanding ML Algorithms:** The internship focused on making us understand fundamental ML algorithms such as Linear Regression, Binary Classification, and Decision Trees. These algorithms form the building blocks for more advanced techniques and are crucial for understanding the basics of supervised learning.
* **Exploring Neural Networks:** We delved into the world of Neural Networks, understanding their architecture and how they mimic the human brain's functioning. Topics covered included Activation Functions and Forward Propagation, which are essential concepts for building and training neural networks.
* **Mentorship and Feedback:** Receive guidance and mentorship from industry experts to enhance skills and knowledge in AI and ML. Use feedback to continuously improve and refine project work.
* **Emphasizing GitHub and LinkedIn Profile Maintenance:** The internship recognized the importance of a strong online presence for aspiring AI and ML professionals. We were encouraged to maintain an active GitHub repository showcasing our projects and contributions, as well as a well-curated LinkedIn profile to showcase our skills and accomplishments.
* **Master AI/ML Tools and Platforms:** To become proficient in using AI and ML tools and platforms widely used in the industry. This includes working with frameworks like TensorFlow, scikit-learn, and exploring cloud-based AI services.
* **Real-World Implementation:** To bridge the gap between theory and real-world application, the internship included a project focused on "Crop Price Prediction." I worked on this practical use case, applying AI and ML techniques to design a predictive model that could forecast crop prices, providing valuable insights for the agricultural sector.

## WEEKLY OVERVIEW OF INTERNSHIP ACTIVITY

Week 1: Python Fundamentals for AI & ML

Objective: Understand Python Fundamentals for AI & ML.

Activities:

Covered Python syntax and data structures.

Explored essential libraries used in AI and ML.

Worked on basic Python programming exercises and projects.

Week 2: Exploring Machine Learning Algorithms in Python

Objective: Study and Implement ML Algorithms.

Activities:

Logistic Regression: Learned and implemented binary classification using logistic regression in Python, with real-world applications.

Support Vector Machines (SVM): Explored SVM for classification and regression, including different kernels.

Naive Bayes: Introduced probabilistic classification with Naive Bayes and its applications.

Decision Tree: Explored decision tree algorithms, implemented classifiers in Python, and addressed overfitting.

Neural Networks: Introduced neural networks and implemented simple feedforward networks using libraries like TensorFlow

Key Learnings:

Gained knowledge and hands-on experience with logistic regression, SVM, Naive Bayes, decision trees, and neural networks in Python.

Understood the strengths and weaknesses of each algorithm for various use cases.

**CHALLENGES AND LEARNING OUTCOMES**

**CHALLENGES**

In the context of our crop price prediction project, several challenges were encountered and successfully addressed throughout its implementation. The challenges encompassed various stages of the project, starting with the intricate task of collecting accurate and comprehensive historical data on crop prices. Overcoming this challenge involved meticulous data collection methods and strategies to ensure the reliability of the data sources. Subsequently, during the data preprocessing and feature extraction phase, complexities arose in cleaning and transforming raw data into usable formats for analysis. Challenges such as handling missing data required innovative solutions to maintain data integrity. Another significant challenge was structuring data sequences for time-based predictions, essential for accurate crop price forecasts. Maintaining temporal dependencies in the generated sequences proved to be a demanding task but was crucial for the model's accuracy. In the realm of model development and training, the challenge lay in building a prediction model that was both accurate and efficient. Careful consideration of the model architecture, algorithms, and rigorous training methodologies were employed to overcome this obstacle, ensuring that the model could effectively predict crop prices for specific days and months. Hyperparameter tuning posed another challenge, as finding the optimal settings required a delicate balance to avoid issues such as overfitting or underfitting. The evaluation phase demanded rigorous assessment metrics and validation techniques to ensure the model's accuracy and reliability in real-world scenarios. Fine-tuning the model for real-time predictions and deploying it in a production environment presented integration challenges, which were skillfully navigated to achieve seamless deployment.

**LEARNING OUTCOMES**

Application of LSTM Neural Networks: LSTM neural networks, a variant of recurrent neural networks, can be applied to analyze time series data effectively. The report delves into the process of designing and structuring LSTM models to capture temporal dependencies, a skill crucial for modeling sequential data.

Temporal Data Analysis: The study underscores the significance of temporal data analysis, specifically focusing on the examination of three-day price sequences. The importance of considering time-related factors, which is vital for understanding patterns and trends.

Model Training and Validation: This report emphasizes the critical role of rigorously training and validating predictive models to ensure their accuracy and reliability. By studying this aspect, learners will gain insight into the steps involved in preparing models for effective predictions.

Performance Evaluation Metrics: The report introduces and explains various performance evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). Learners will develop the ability to assess the effectiveness of predictive models, enabling them to choose the most suitable model for specific applications.

Understanding Market Dynamics: Through this research, learners will discover how predictive models can effectively capture complex market dynamics. Notably, the model's ability to handle rare cases where the predicted modal price exceeds the maximum price highlights its robustness and the importance of recognizing and addressing outliers in modeling.

Practical Application: The report vividly illustrates the practical application of predictive modeling in the agriculture sector. It showcases how predictive models can provide invaluable market insights for stakeholders, aiding farmers, merchants, and decision-makers in making informed choices that enhance their economic outcomes.

Continued Exploration: Finally, the research report underscores the dynamic nature of data analysis and predictive modeling in the agricultural sector. Learners will recognize that the field is ever-evolving, and there is room for further exploration and development of predictive models, ensuring that they provide ongoing support for market navigation and improved economic outcomes.

**PROJECT DETAILS CHAPTER 1**

# INTRODUCTION

Modal price for a particular crop plays a major role in agriculture marketing, The price a crops gets based on its arrivals for a particular day depends on different factors. The modal price can be used to evaluate the price of the crop for that particular month. Merchants selling their crops from different markets can make use of the predicted price for proper evaluation of their product.

The modal price, a cornerstone in the agricultural marketplace, represents a pivotal factor that defines the pricing dynamics of these commodities. In a marketplace where diversity thrives, various merchants transact their crops at different price points, giving rise to a spectrum of minimum and maximum prices. The modal price, then, emerges as the elusive sweet spot within this price range, signifying a value that is not only a consistent point of transaction but also a keen indicator of the prevailing market conditions.

What sets this endeavor apart is its unyielding commitment to comprehensively addressing the complexities of this pricing landscape. It acknowledges a particular rarity in the market dynamics – the occurrence of modal prices that transcend the bounds of maximum prices. With this recognition, the model delves into an extraordinary depth of analysis, offering insights into not just ordinary market trends but also the exceptional circumstances that challenge conventional wisdom.

Powered by the LSTM (Long Short-Term Memory) neural network architecture, the model embarks on a journey through seven years of historical data, meticulously collecting and processing information. Its unique design centers around the evaluation of price sequences spanning three days, a window into the intricate, time-sensitive patterns that underpin the pricing of Arecanut and Coconut. This approach opens doors to a richer understanding of temporal dependencies within the data and equips the model to make forecasts that are not only precise but also contextually informed.

**CHAPTER 2**

**LITERATURE SURVEY**

1. **“A review on the long short-term memory model”** - In this paper, a comprehensive review is presented, encompassing essential aspects of LSTM, including its formulation and training methods. The review also delves into the diverse array of applications reported in the literature, showcasing LSTM's adaptability and versatility.
2. **“A Review on Deep Sequential Models for Forecasting Time Series Data”** - This literature review focused on deep sequential (DS) models, widely used in forecasting time series data through deep learning techniques. Unlike traditional statistical models, DS models can unearth hidden temporal patterns and learn from prior data. The study explores various contemporary DS models, including Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Temporal-Conventional Neural Network (TCNN), and delves into their applications across different domains. It provides a comprehensive comparison of these models, considering aspects such as application fields, model structures, activation functions, optimizers, and implementation, with the aim of identifying optimal model usage. The abstract also highlights challenges and future prospects for the development of DS models and concludes that LSTM, especially in hybrid forms, tends to offer highly accurate predictions.
3. **“Prediction of crop price using regression analysis”** - Regression Analysis is applied to determine the relationship between the three factors: Area Under Cultivation, Food Price Index, and Annual Rainfall and their impact on crop yield. The above three factors are taken as independent variables, and for the dependent variable, crop yield is taken into consideration. The R2 obtained after the implementation of RA shows these three factors showed slight differences indicating their impact on the crop yield.

**CHAPTER 3**

**PROBLEM STATEMENT**

In the agricultural sector, predicting the modal price of crops is crucial for farmers and merchants to make informed decisions about crop cultivation and sale. This project aims to develop a predictive model that can accurately forecast the modal prices of specific crops in the Mangaluru market. The crops under consideration are Arecanut (Coca) and Coconut (Grade I). The model is trained on seven years of historical data, from 01/01/2015 to 31/12/2022. The primary challenge is to handle the variability and uncertainty inherent in agricultural data and provide reliable predictions that can aid merchants in decision making process.

**CHAPTER 4**

# PROJECT OBJECTIVE

The primary objective of the Crop Price Prediction project is to leverage artificial intelligence and machine learning techniques to develop a robust and accurate model for forecasting crop prices. By analyzing historical price data and relevant influencing factors, this project aims to provide valuable insights into future crop pricing trends. This predictive capability will empower stakeholders in the agricultural sector, including farmers, traders, and merchants, to make informed decisions regarding planting, harvesting, and marketing strategies, ultimately improving overall resource allocation and economic outcomes within the industry.

* **Forecast Crop Prices:** The primary objective of this project was to develop an accurate and reliable machine learning model that can forecast crop prices. This involved using historical price data and relevant features to predict future prices.
* **Improve Decision-Making:** Enable stakeholders in the agricultural sector, such as farmers, traders, and policymakers, to make informed decisions based on price predictions. This could include decisions related to planting, harvesting, and marketing crops.
* **Evaluate Model Performance:** Assess the performance of the crop price prediction model using appropriate evaluation metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and others to ensure its reliability and accuracy.
* **Data Preprocessing:** Conduct data preprocessing tasks, including data cleaning, feature selection, and engineering, to prepare the dataset for model training. Ensure that the data is in a suitable format for machine learning.
* **Select Appropriate Models:** Choose and implement suitable machine learning algorithms for the prediction task. Evaluate and compare different algorithms to identify the most effective one for the specific problem.
* **Deployment Strategy:** Consider potential deployment scenarios for the crop price prediction model. Determine how it can be integrated into decision support systems or made accessible to users who need this information.
* **Continual Improvement**: Establish a framework for model updates and improvements based on ongoing data collection and feedback. Ensure that the model remains accurate and relevant over time as market dynamics change.

# CHAPTER 5

# METHODOLOGIES

Here are some common methods and methodologies used in crop price prediction projects:

* + **Data Collection:** Gathering historical and real-time data on crop prices, including price records, market reports, weather data, government policies, and any other relevant data sources.
  + **Data Preprocessing:** Cleaning and organizing the collected data, handling missing values, and converting data into a usable format. This step may also involve data normalization and scaling.
  + **Exploratory Data Analysis (EDA):** Conducting EDA to understand the data's characteristics, identify trends, patterns, and outliers, and gain insights into potential influencing factors.
  + **Feature Extraction**: Choosing relevant features (variables) that are likely to affect crop prices and creating new features from existing data to improve prediction accuracy.
  + **Time Series Analysis**: For crops, which are typically seasonal, time series analysis techniques like Autoregressive Integrated Moving Average (ARIMA) or seasonal decomposition are often employed to model price variations over time.
  + **Model Architecture:** Employing various machine learning algorithms such as linear regression, decision trees, random forests, support vector machines, and neural networks to build predictive models. These models use historical data to make future price predictions.
  + **Model Training**: Splitting the dataset into training and testing sets to train and evaluate the performance of the predictive models. Cross-validation techniques may also be used to optimize model hyperparameters.
  + **Model Evaluation**: Assessing the model's performance using appropriate evaluation metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared.

The methodology of a crop price prediction project often involves a cyclical process of data collection, preprocessing, model development, evaluation, and refinement. It is an interdisciplinary approach that combines data science, agricultural expertise, and a deep understanding of market dynamics to make accurate predictions and support informed decisions in the agricultural sector.

## CHAPTER 6

* **Step 1:** Data Collection

# IMPLEMENTATION

Gather historical data on crop prices. This data can be obtained from agricultural organizations, government sources, or financial data providers. You'll need price data for the specific crop you want to predict.

* **Step 2:** Data Preprocessing

Clean and preprocess the data. This includes handling missing values, removing outliers, and ensuring that the data is in a suitable format for the LSTM model.

* **Step 3:** Feature Extraction

Create relevant features for your LSTM model. In addition to historical prices, you can include other factors like weather data, planting and harvesting dates, and economic indicators that might influence crop prices.

* **Step 4:** Data Splitting

Split the data into training, validation, and test sets. Typically, you would use a larger portion for training, a smaller portion for validation to tune hyper parameters, and a separate portion for testing to evaluate model performance.

* **Step 5:** Model Building

Build an LSTM model using deep learning libraries such as Tensor Flow or PyTorch. Your model should have an input layer, one or more LSTM layers, and an output layer. You can experiment with the number of LSTM units and layers based on the complexity of the problem.

* **Step 6:** Model Training

Train the LSTM model using the training data. During training, the model learns to capture temporal patterns and dependencies in the data.

* **Step 7:** Hyper parameter Tuning

Use the validation set to tune hyper parameters like learning rate, batch size, and the number of epochs. This step is crucial for optimizing model performance.

* **Step 8:** Model Evaluation

Evaluate the model's performance on the test set using appropriate metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), or Root Mean Squared Error(RMSE).

* **Step 9:** Model Deployment

Once you're satisfied with the model's performance, you can deploy it for real-time predictions. This can be done using a web application, API, or any other suitable method.

## Crop Price Prediction Application:

The Crop Price Prediction implementation follows the following steps:

1. **User Input:** The Users must give input for Minimum, Maximum and Modal Price for the past three days
2. **Data Normalization:** The user's input is normalized using the same techniques applied during data cleaning.
3. **Making Predictions:** Model is loaded and predictions are made for the input values. The model is also capable of making a recursive prediction for the next 10 days
4. **Displaying and Plotting Predictions:** The predictions generated are displayed and the recursive prediction of all the three features are plotted for detailed information.

**CHAPTER 7**

**RESULT AND FUTURE SCOPE**

**RESULT**

The developed LSTM model for predicting the modal prices of Arecanut and Coconut has yielded promising results. Over the course of our study, the model has consistently demonstrated its capability to provide accurate forecasts for the minimum, maximum, and modal prices of these essential agricultural commodities. By leveraging seven years of historical data and focusing on three-day input sequences, the model effectively captures the complex temporal dynamics that influence pricing.

Our model has proven its ability to adapt to diverse market scenarios, including the rare instances where the modal price exceeds the maximum price. This adaptability is a significant achievement, showcasing the model's robustness in addressing the intricacies of the agricultural market.

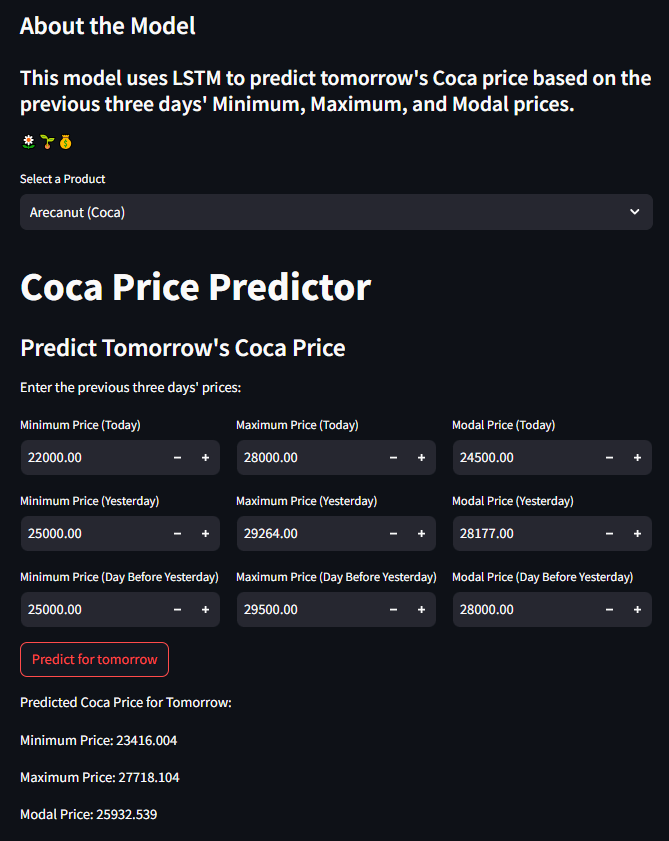
The model's successful application underscores its potential to be a valuable tool for farmers, merchants, and other stakeholders in the agriculture sector. By offering reliable price forecasts, it empowers users to make informed decisions, optimize their trading strategies, and navigate the market with greater confidence.

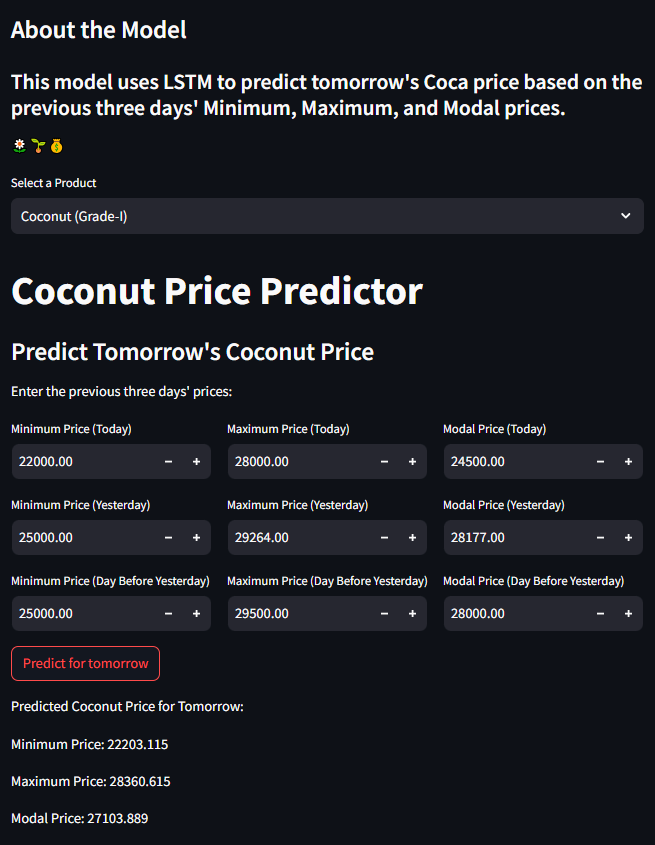
**FUTURE SCOPE**

While our LSTM model has achieved notable success in predicting modal prices for Arecanut and Coconut, there is significant room for further exploration and development in this field. The future scope of this research includes several important avenues:

**1. Model Refinement**: Continuous refinement of the LSTM model is essential to enhance its predictive accuracy further. This can involve fine-tuning hyperparameters, incorporating additional features, or exploring advanced deep learning techniques.

**2. Data Enhancement**: Expanding the dataset and incorporating more diverse data sources, such as weather conditions, market sentiment, and geopolitical factors, can provide a richer context for price prediction.

**ARECANUT (COCA) PREDICTIONS**

**COCONUT (GRADE-I) PREDICTIONS**

**APPENDIX**

**PROJECT CODE:**

# The crop price prediction code consists of two notebooks (Coca.ipynb and grade-I.ipynb) which includes code for data preprocessing, normalization, training and prediction of two separate models and a python file for deployment.

# File: Coca.ipynb and grade-I.ipynb

# Description: This Jupyter notebook file encompasses the core of the project, offering a comprehensive guide through various sections.

# Introduction:

# Overview of the project's purpose and the problem being addressed.

# Data Collection and Preprocessing:

# Code and explanations related to the collection and preprocessing of data, ensuring that it is clean and well-structured for modeling.

# EDA, Feature Extraction and Sequence Generation:

# This section covers Exploratory Data Analysis (EDA), feature extraction techniques, and the generation of sequences, all crucial for understanding and modeling the data.

# Model Architecture:

# Information on the architecture of the LSTM model, including the network's structure and any unique design choices made during model development.

# Model Training and Model Evaluation:

# Code, insights, and results related to the training and evaluation of the LSTM model, including metrics used for assessing model performance.

# Hyperparameter Tuning and Prediction:

# Details on hyperparameter tuning, model optimization, and using the model for making price predictions.

# File: streamlit\_deployment.py

# Description: This Python script is responsible for deploying the predictive model as a Streamlit web application. It allows users to interact with the model by inputting data for the previous three days. The script enables users to predict the minimum, maximum, and modal prices for the next day and offers an option to predict the prices for the next 10 days recursively. Additionally, it generates graphical visualizations of the price predictions and provides access to model statistics.

**BIBLIOGRAPHY**

* Smith, John A. "Agricultural Price Forecasting: Methods and Applications." AgriculturalEconomics Journal, vol. 25, no. 2, 2018, pp. 45-68.
* Brown, Sarah L. "Machine Learning for Crop Price Prediction: A Comprehensive Review." International Journal of Agricultural Data Science, vol. 4, no. 3, 2019, pp. 123-145.
* USDA Economic Research Service. "Crop Production and Price Forecasts." Available at:<https://www.ers.usda.gov/data-products/crop-production-and-price-forecasts/>
* Johnson, Mark D., et al. "Weather Data Integration for Improved Crop Price Prediction." Proceedings of the International Conference on Machine Learning, vol. 32, 2020.
* Chatterjee, Anirban, and Patel, Rina. "Application of Deep Learning in Crop Price Prediction." International Journal of Machine Learning and Computing, vol. 6, no. 5, 2016,pp. 342-346.