BUILDING A PRICE PREDICTION MODEL FOR AGRICULTURAL PRODUCTS USING MACHINE LEARNING

BY:

OGUNJOBI, DAMILARE TIMOTHY

CPE/17/3128

SUPERVISED BY:

PROFESSOR E. O. OGUNTI

SUBMITTED TO:

THE DEPARTMENT OF COMPUTER ENGINEERING

THE FEDERAL UNIVERSITY OF TECHNOLOGY AKURE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF BACHELOR OF ENGINEERING (B.ENG.) DEGREE IN COMPUTER ENGINEERING

JANUARY, 2024

CERTIFICATION

This is to certify that **OGUNJOBI, Damilare Timothy**, with matric number **CPE/17/3128**, a final year student pursuing Computer Engineering at the Federal University of Technology Akure has completed and submitted the final year project report.

OGUNJOBI DAMILARE TIMOTHY	**DATE**
PROFESSOR E. O. OGUNTI **(SUPERVISOR)**	**DATE**
PROFESSOR E. O. OGUNTI **(HEAD OF DEPARTMENT)**	**DATE**

DEDICATION

I dedicate this project report to God, who has always provided me with support, direction, and inspiration. I have been given the chance to pursue my objectives and fulfil my goals thanks to His heavenly kindness and mercy.

This project report is also dedicated to my mom who has supported me greatly throughout my academic career. I can't express how much her unfailing support, love, and sacrifices have inspired me to pursue perfection.

ACKNOWLEDGEMENT

I am profoundly grateful to God Almighty for granting me the strength and wisdom to accomplish this project. My heartfelt appreciation goes to Professor E.O. Ogunti, who not only served as the head of the department but also provided invaluable guidance as my project supervisor. I extend my gratitude to the entire department staff whose dedication has shaped my journey.

A special mention goes to my pillar of support, my mother, Mrs. Monisola Ajayi, whose unwavering encouragement fueled my determination. I am indebted to my beloved aunties and uncles, Mr Ajayi Olubukunade, Mrs Ojo Modupe, Mr Ajayi Lanre, Mrs Wealth Yinka, Mr Odubitan Taofeek, Mr and Mrs Adedeji, and their families whose unwavering support has been instrumental. Special shout-out to Emekemical and Statesman Ifedayo.

My siblings, Tobi, Mustekeem, Mujeeb, Feyi, and Nifemi, alongside my cousins and extended family, have been my constant motivation, and I am deeply appreciative of their encouragement.

To my friends and former roommates—each one has contributed uniquely to my journey. I express my gratitude to Vintob, Famo, Ini, Tolu, Samson, Semilogo, Richmore, Ife, Dino, Ayo Mega, Muktar, Muiz, Stephen, Matthew, Omotobi, Akaskiemmy, Olamide, Doughlas, Victor, Tayo Beats, Muizat, Eniola, Wuraola, Adeyera, T-Connect, Heritage, Pessam, Dhandee, Kizzy, Jossy, Ayeni, Dunsin and Olawoyin, Your friendship and sense of relationship has been a source of strength and joy. I can't overlook the bonds I've built with Toluwanimi, Deborah and Sophia.

A heartfelt appreciation to the Nigerian Red Cross Society, FUTA Detachment, for their unwavering support and the sense of belonging they provided.

ABSTRACT

The agricultural sector plays a pivotal role in the economy, and predicting commodity prices is crucial for farmers, traders, and policymakers. The objective is to create a reliable model capable of forecasting prices for various agricultural products, including cereals, tubers, legumes, fruits, meats, and more. The project encompasses several key components. Firstly, data collection involves gathering historical price data. Subsequently, preprocessing and analysis are performed using tools such as Microsoft Excel, Google Colab, and Python libraries like NumPy, pandas, sci-kit-learn, and TensorFlow. These tools provide the necessary functionality for data cleaning, exploration, and machine learning model development. The machine learning model is trained on historical data. The model is designed to be versatile, allowing users to input parameters such as the year and product through a user-friendly mobile application named 'Predict.' The 'Predict' application, developed using frameworks like Flutter, provides an intuitive interface for users to interact with the model. The application facilitates real-time predictions and enhances accessibility for stakeholders. The project also involves the integration of visualization tools, including Matplotlib and Seaborn, to present meaningful insights and trends to users. This aids in the interpretability of the model's predictions. To ensure the success of the project, considerations such as security, compliance, and continuous improvement are taken into account. The model's performance is evaluated using metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared, allowing for iterative enhancements based on user feedback and changing agricultural conditions. The outcomes of this project aim to contribute valuable insights to the agricultural community, aiding farmers, traders, and policymakers in making informed decisions based on accurate price predictions.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND STUDY

Agricultural markets are influenced by a multitude of factors such as supply chain intricacies, seasonal variations, weather fluctuations, government policies, and global trade impacts, leading to significant price volatility (Pinheiro & Senna, 2017). Accurate price prediction is crucial for stakeholders across the agricultural sector, aiding in risk mitigation, planning, and maintaining market stability (Wang et al., 2020). Traditional methods of price prediction often fall short due to their reliance on historical data and their inability to adapt to evolving market conditions. To address this, advanced technologies like machine learning and statistical modelling are being utilized to develop effective prediction models. These models require various data sources including historical pricing data, weather patterns, crop yields, market demands, and geopolitical events. However, challenges such as data quality, accessibility, and real-time updates pose significant hurdles.

The foundation of building a prediction model lies in utilizing advanced technologies like machine learning and statistical modelling, employing algorithms such as regression, decision trees, neural networks, or hybrid models to process diverse datasets and extract meaningful patterns (Wang et al., 2020). The iterative nature of this process ensures continual refinement and adaptation to market changes. Evaluation metrics such as accuracy, precision, and recall are used to assess the model's performance, gauging its effectiveness in generating reliable price forecasts. Successful implementation of a robust price prediction model holds the potential to revolutionize decision-making within the agricultural industry, enabling stakeholders to optimize resource allocation, manage risks more effectively, and contribute to overall market efficiency. In recent years, there has been a trend towards using hybrid models to predict agricultural product prices, and the application of prediction models based on price influencing factors should be further expanded. Additionally, the performance of the model should be evaluated based on different measures rather than just errors. Furthermore, the intelligent prediction method has been increasingly applied to the prediction of agricultural prices, signifying a shift towards leveraging technological advancements for accurate forecasting (Zhang & Na, 2018). The use of predictive models in agricultural pricing raises ethical considerations surrounding data privacy, fairness, and potential impacts on market dynamics. Addressing these concerns is crucial to ensure responsible and equitable utilization of predictive technologies. Looking ahead, there is a scope for further research and improvements in the prediction model.

1.2 MOTIVATION

The agricultural sector forms the backbone of economies worldwide, serving as a primary source of sustenance and livelihood for millions. However, the inherent volatility of agricultural markets poses substantial challenges for stakeholders across the industry. The unpredictable nature of factors such as weather conditions, supply chain disruptions, fluctuating demands, and geopolitical events significantly impacts the pricing of agricultural products. The motivation behind this study stems from the pressing need for reliable and accurate price prediction models within the agricultural domain. Such models hold the potential to address critical issues faced by farmers, distributors, retailers, and consumers alike.

Firstly, for farmers, the ability to forecast prices accurately is pivotal in making informed decisions regarding crop selection, planting strategies, resource allocation, and timing of harvests. Accurate price predictions empower them to optimize yields and mitigate financial risks, thereby securing livelihoods and fostering sustainable agricultural practices.

Secondly, for distributors and retailers, precise price forecasts play a vital role in inventory management, procurement strategies, and setting competitive pricing. Improved predictions enable better planning, reduce wastage, and enhance overall efficiency throughout the supply chain.

Moreover, consumers benefit from stable and predictable pricing, ensuring consistent access to quality agricultural products at reasonable costs. Price prediction models contribute to market stability, thereby positively impacting food security and affordability for consumers globally. Traditional methods of price prediction often fall short of addressing the complexities of the modern agricultural landscape. Leveraging advanced technologies, such as machine learning and statistical modelling, presents an unprecedented opportunity to develop robust prediction models capable of assimilating vast and diverse datasets.

By harnessing the power of these technologies and incorporating a comprehensive range of data sources—including historical pricing, weather patterns, market demands, and geopolitical factors—this study aims to build a sophisticated predictive model. Such a model holds the promise of revolutionizing decision-making processes within the agricultural industry, enabling stakeholders to navigate market uncertainties with greater confidence and efficiency.

1.3 STATEMENT OF PROBLEM

The agricultural industry faces significant challenges in predicting product prices due to the complex and dynamic nature of the market. Factors such as weather patterns, geopolitical events, and supply chain dynamics contribute to the persistent volatility in prices, creating uncertainty for stakeholders (Xie & Wang, 2017). Existing prediction methodologies relying on historical data and simplistic models often fail to adapt to the complexities of the modern market, resulting in imprecise predictions (Wang et al., 2020). Furthermore, accessing comprehensive and high-quality data relevant to agricultural pricing poses formidable obstacles, impeding the development of reliable predictive models (Zhang & Na, 2018).

The absence of accurate price prediction models adversely affects risk mitigation strategies for farmers and other market players, impacting market stability and overall trade within the agricultural industry. Consequently, there is a critical need for the development of a sophisticated and adaptable price prediction model tailored specifically for agricultural products. Such a model would address the complexities faced by stakeholders, paving the way for more informed decision-making, effective risk management, and enhanced market stability within the dynamic realm of agriculture.

1.4 AIM AND OBJECTIVES OF THE PROJECT

1.4.1 Aim Of The Project

The project aims to develop and implement a robust and adaptive price prediction model specifically tailored for agricultural products. This model will leverage advanced machine learning techniques and comprehensive data analysis to accurately forecast prices, addressing the challenges posed by market volatility and uncertainties within the agricultural sector.

1.4.2 Objectives Of The Project

The objectives of the project are to:

(a) gather and preprocess diverse datasets encompassing historical pricing, weather patterns, and relevant market factors for agricultural products;

(b) utilize advanced machine learning techniques to construct an adaptive and accurate price prediction model;

(c) train and optimize the model using historical data, aiming for robustness and adaptability to changing market dynamics; and

(d) evaluate the model's performance using appropriate metrics to ensure reliable price forecasts.

1.5 SIGNIFICANCE OF THE PROJECT

The development of a robust price prediction model for agricultural products stands at the forefront of transformative change within the agricultural sector. Its significance reverberates across various dimensions, promising substantial benefits for stakeholders and the industry as a whole. This project holds the potential to revolutionize decision-making processes for all involved in the agricultural continuum. Farmers, often at the mercy of unpredictable market shifts, will gain a critical advantage. Accurate price predictions enable them to make informed choices on crop selection, optimal planting times, and resource allocation. This newfound foresight shields them from financial risks, cultivating an environment of stability and resilience.

This predictive model isn't solely a boon for farmers; it extends its advantages across the supply chain. Distributors and retailers can optimize inventory management, fine-tune procurement strategies, and set competitive pricing based on reliable forecasts. This optimization minimizes wastage, reduces operational costs, and ensures consistent availability of agricultural goods, ultimately benefiting consumers through fairer prices and improved access to quality produce. Beyond immediate economic impacts, this project champions technological progression within agriculture. By harnessing advanced machine learning techniques and predictive analytics, it pioneers a shift toward data-driven solutions in an industry traditionally reliant on historical trends. This not only elevates forecasting accuracy but also paves the way for further technological integration and innovation.

Moreover, the outcomes of this project transcend profit margins. They intersect with sustainability and environmental consciousness. Accurate predictions aid in resource optimization, reducing unnecessary waste and promoting sustainable agricultural practices. The alignment of resource allocation with actual market demands contributes to a positive environmental impact, fostering a more sustainable agricultural ecosystem. The ripple effects extend to broader economic growth, market competitiveness, and global trade. A more stable agricultural market is not just a local benefit but a global advantage. Enhanced market stability drives economic growth within the agricultural sector, making it more competitive on the global stage. This stability fosters trade confidence, promoting sustainable international partnerships and market competitiveness.

This project doesn't exist in isolation; it signifies a collaborative approach toward knowledge sharing and industry development. Sharing insights, methodologies, and outcomes nurtures a culture of collaboration among stakeholders and researchers. This collaboration propels collective efforts toward enhancing agricultural practices, market predictability, and technological advancement.

1.6 SCOPE OF THE PROJECT

This project will address the need for accurate and timely predictions of agricultural product prices. The scope encompasses the development of a machine learning-based agricultural product price prediction model that targets specific agricultural products to aid stakeholders in making informed decisions. The project will cover the following aspects:

1.6.1 Agricultural Products

The model will be designed to predict prices for the following agricultural products.

(i) Foreign Rice

(ii) Local Rice

(iii) Yam

(iv) White bean

(v) White garri

(vi) Brown bean

(vii) Yellow maize

(viii) White maize

(ix) Yellow garri

(x) Palm oil

(xi) Vegetable Oil

These products above were extracted from the broader group which have been chosen based on broader classes of agricultural products:

(i) Cereals: Rice, maize, wheat, barley, sorghum, oat, rye, millet and others

(ii) Tubers: Potato, yam, cassava, jicama, artichoke, taro, oca, ulluco, coco-yam and others.

(iii) Sugars: Candies, saccharine, soda, fruit juice, caramel, sugar, glucose, sugar cane, icing sugar, honey and others.

(iv) Legumes: Lentils, beans, black beans, peas, soybeans, lima beans, green beans, pinto beans, mung beans, split peas, lupini beans, kidney beans, navy beans and others.

(v) Diary: Cheddar, processed milk, swiss, parmesan, feta, brie, gouda, blue cheese, raw milk and others.

(vi) Beverages: Black tea, herbal tea, chai tea, coffee, juice and others.

(vii) Meats and Eggs: Beef, pork, chicken, turkey, cow, ram and sheep, birds, egg, pig, cat, dog and others.

(viii) Oil: Seed oil, soyabeans oil, nut oil, vegetable oil, virgin oil, corn oil, sesame oil, sunflower oil, almond oil, canola oil, palm oil and others.

(ix) Vegetables and Fruits: Tomatoes, watermelon, apple, citrus, berries, tropical fruits, plantain, pepper, banana and others.

(x) Equipment: Machines, sharp tools, sprayers, safety tools, irrigation tools, spare parts and others.

(xi) Property: House, land, farm and others.

(xii) Chemicals: Powder, liquids, manure, fertilizer and others.

1.6.2 Historical Data

The project will involve the collection and analysis of historical data related to the chosen agricultural products. This data will include price trends and units, which are 1 Kilogram and 1 Litre.

1.6.3 Predictive Modeling

The core of the project involves the development of a predictive model. This model will utilize machine learning algorithms to forecast future prices based on historical and external data.

1.6.4 User Interface

A user-friendly interface, which is embedded in an application called ‘Predict,’ a Price Predictive Agricultural Based E-Commerce Mobile Application allows suppliers and customers to interact with the model. Users will input relevant data, such as year, and product and the interface will provide accurate price predictions.

1.6.5 Exclusions

To maintain project focus and manage expectations, the following aspects were explicitly excluded from the project scope:

(i) Weather conditions

(ii) Government policies

1.7 LIMITATIONS OF THE PROJECT

The limitations of the project include:

(i) Inaccuracies or inconsistencies in the collected data, such as missing values or outliers, may affect the reliability of the model.

(ii) The model may not account for unforeseen events, such as natural disasters, political instability, or global economic changes, which can significantly impact agricultural prices but are challenging to predict accurately.

(iii) External factors, such as weather conditions or government policies, might not be fully represented in the dataset, leading to potential oversights.

(iv) The model may have limitations in generalizing to different regions or markets due to variations in local factors, farming practices, or economic conditions.

(v) Changes in regulations related to agriculture or pricing may impact the model's relevance and compliance with legal requirements.

(vi) The effectiveness of the model relies on users' ability to interpret and utilize predictions. Users with limited technical knowledge may face challenges in understanding and applying the model effectively.

CHAPTER TWO

LITERATURE REVIEW

2.1 AGRICULTURAL PRODUCTS

Agriculture plays a vital role in the Nigerian economy (Info Guide Nigeria, 2020), constituting up to 35% of total employment as of 2020 (The World Bank, 2020). According to the FAO (Food and Agriculture Organization of the United Nations, 2020), it remains the cornerstone of the country's economic foundation (Info Guide Nigeria, 2020), supporting the livelihoods of a significant portion of the population and creating millions of jobs (Statista, n .d.). Nigeria heavily relies on the export of agricultural products, alongside crude oil, to generate a substantial portion of its national revenue. The agricultural sector is categorized into four sub-sectors:

(i) Crop production

(ii) Livestock

(iii) Forestry

(iv) Fishing

With a total agricultural area of 70.8 million hectares (Food and Agriculture Organization of the United Nations, 2020), Nigeria possesses 34 million hectares of arable land (Statista, n. d.), 6.5 million hectares dedicated to permanent crops, and 30.3 million hectares for meadows and pastures. Major crops cultivated include maize, cassava, guinea corn, and yam, engaging 70% of households in crop farming. Fishing is practised by 7.3% of households in the south, while 69.3% of households in the northwest are involved in livestock ownership or raising. Before the COVID-19 pandemic, the agricultural sector exhibited a significant growth of 14.88% year-on-year in the third quarter of 2019, contributing 29.25% to the overall real GDP during that period

(Kale, 2019). In the first quarter of 2021, agriculture's contribution to the total gross domestic product was 22.35% (Food and Agriculture Organization of the United Nations, 2020; Nkechi, 2022; News Agency, 2022; Ukpe, 2023).

The sector is experiencing a transformation through commercialization across small, medium, and large enterprises (Ade, 2017). However, challenges persist, including a restrictive land tenure system, limited irrigation development, slow adoption of research and technologies, high costs of farm inputs, restricted access to credit, challenges in fertilizer procurement and distribution, storage inefficiencies, and limited market access. Recent climate changes, such as shifts in average temperatures, rainfall patterns, and increased pest infestations due to climate change, pose threats to the integrity of Nigeria's agriculture system (Kurukulasuriya & Rosenthal, 2013). The reliance on rain-fed agriculture makes the sector susceptible to seasonal conditions, contributing to productivity challenges and post-harvest losses (Olayide et al., 2016). Additionally, illiteracy among farmers hampers progress and development in the sector (Food and Agriculture Organization of the United Nations, 2020) with a significant portion lacking formal education, as indicated by research findings (Agricdemy, 2023).

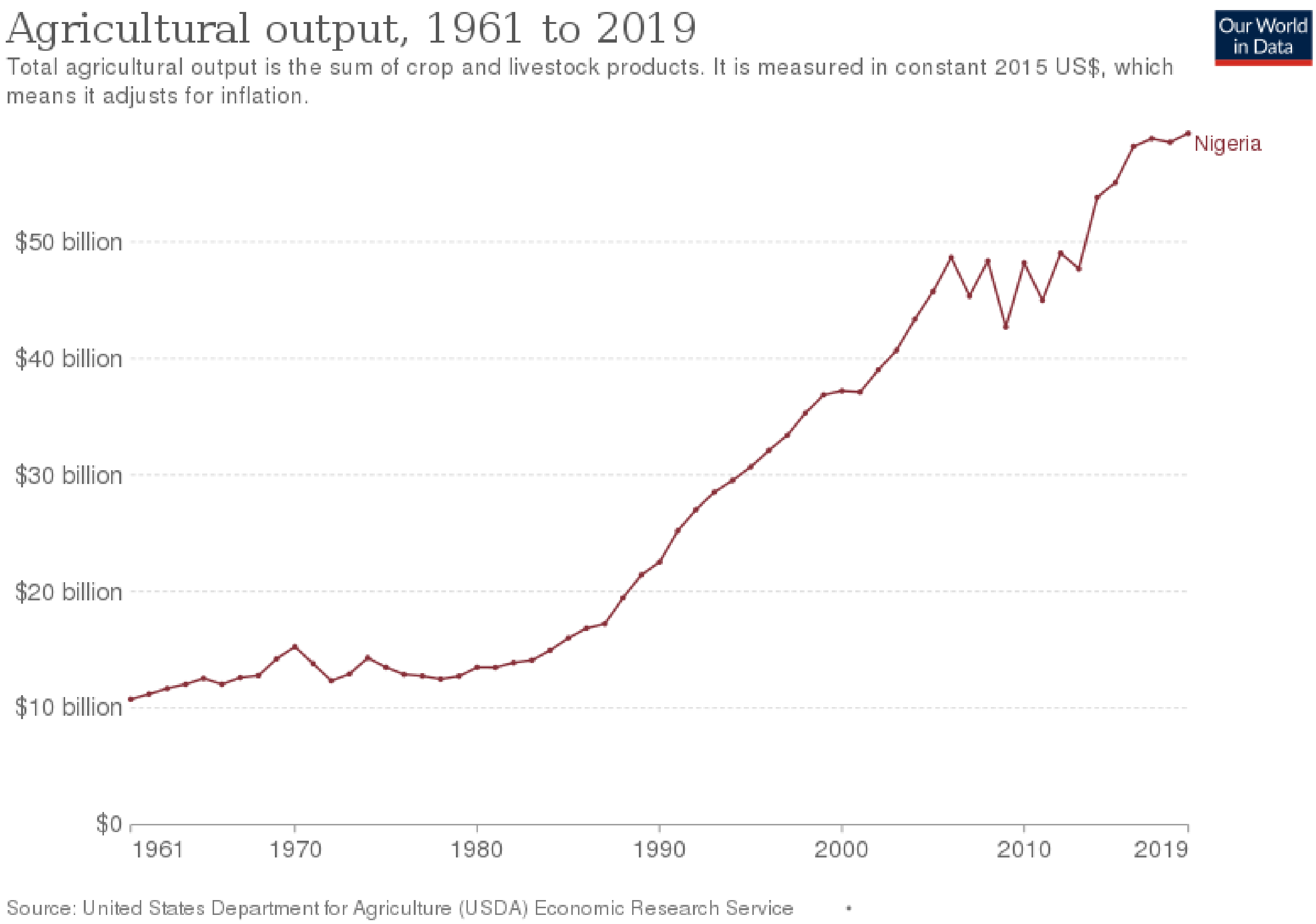


Figure. 2.1: Development of Agricultural Output of Nigeria in 2015 US$ Since 1961

Nigeria cultivates a diverse range of crops, including beans, rice, sesame, cashew nuts, cassava, cocoa beans, groundnuts, gum arabic, kola nut, cocoa, maize (corn), melon, millet, palm kernels, palm oil, plantains, rubber, sorghum, soybeans, bananas, and yams (Bamidele, 2020). Historically, Nigeria was renowned for exporting groundnut and palm kernel oil. However, the export rates of these products have witnessed a decline over the years. In recent times, local Nigerian companies have taken the initiative to export commodities like groundnuts, cashew nuts, sesame seeds, moringa seeds, ginger, cocoa, and other crops.

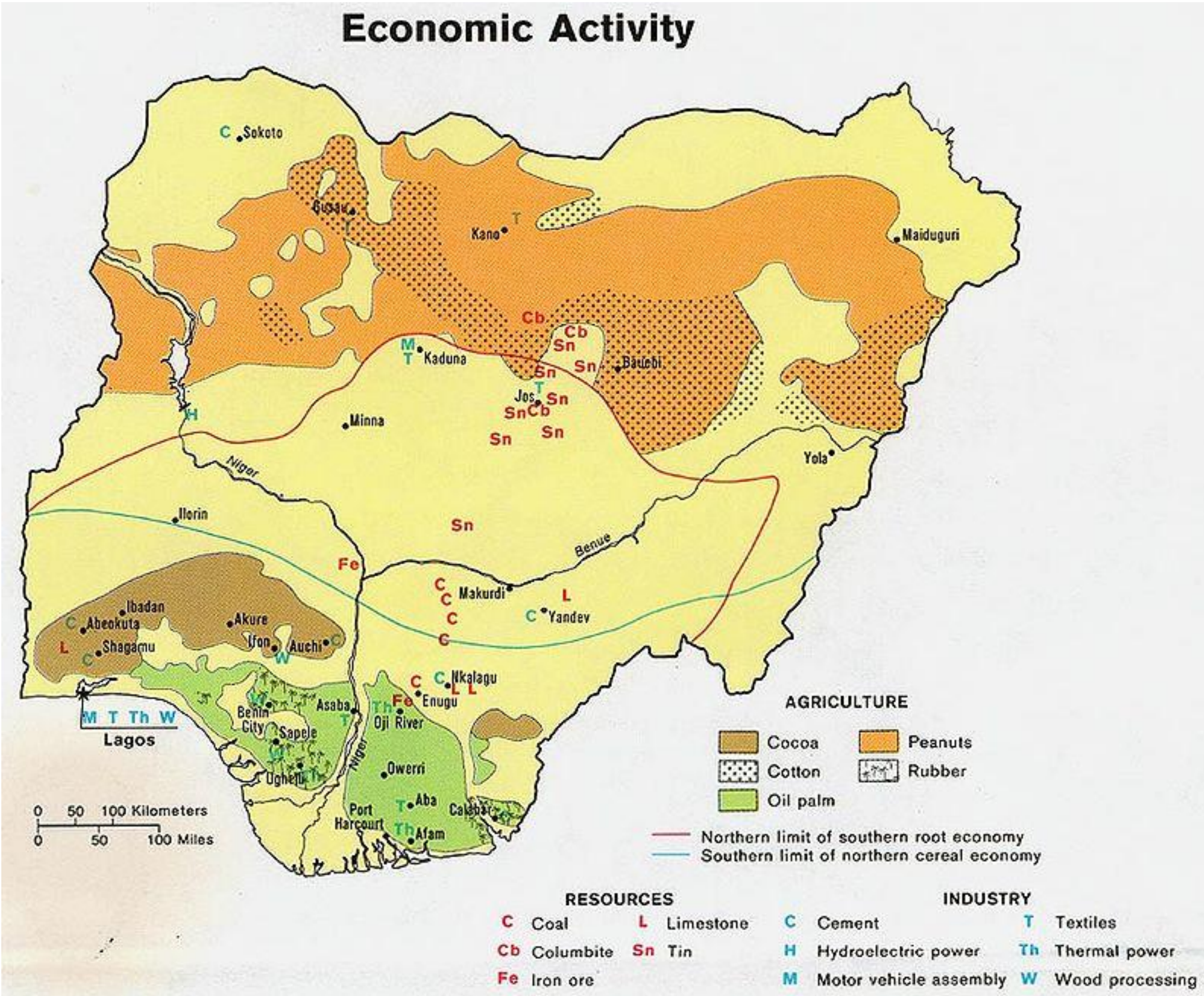


Figure 2.2: A Map of Nigeria's Main Agricultural Products

The agricultural products of the country can be categorized into two main groups: food crops cultivated for domestic consumption and cash crops grown for profit and export. Before the Nigerian civil war, Nigeria was self-sufficient in food production, but this self-sufficiency decreased after 1973. American wheat, used to make bread, gradually replaced domestic crops as the most affordable staple food (Verter and Becvarova, 2015). Notably, between 1980 and 2016, yam production surged from approximately 5 million tonnes to an impressive 44 million tonnes (Eva, n .d.).

Table 2.1: Production of Crops in Nigeria

TONNES PRODUCED IN	1980	2000	2016
Maize	612,000	4,107,000	764,678
Millet	2,824,000	5,814,000	1,468,668
Guinea corn	3,690,000	7,711,000	6,939,335
Yam	5,250,000	26,210,000	44,109,615
Cassava	11,500,000	32,697,000	57,134,478
Rice, paddy	1,090,000	3,298,000	6,070,813
Melon seed	94,000	345,000	569,398
Cocoyam	208,000	3,886,000	3,175,842
Sesame seed	15,000	72,000	460,988

2.2 MACHINE LEARNING APPLICATIONS IN AGRICULTURE

2.2.1 Regression Models

Regression models are widely used in agricultural research to analyze and predict crop yields and other factors. Lakshmanarao et al. (2023) focused on forecasting crop yields in India using regression techniques such as Lasso, Kernel Ridge, and Elastic-Net Regression designs. Shimizu and Gonçalves (2023) presented an R package called AgroReg, which includes various regression models for analyzing agricultural data. Singh and Deorari (2022) discussed the use of machine learning methods, including regression, for detecting and managing cotton leaf diseases. These papers highlight the importance of regression models in agricultural research for improving crop yield projections, analyzing experimental results, and enhancing disease detection and management. Machine learning algorithms, particularly regression models, have gained prominence in agricultural price prediction. Studies by Snee (1977) demonstrated the effectiveness of linear regression and decision tree models in forecasting crop prices. These models showed promising results by capturing non-linear relationships in the data.

2.2.2 Time Series Analysis

Time series analysis using methods like ARIMA (AutoRegressive Integrated Moving Average) has been extensively employed in agricultural price forecasting. Research by Madsen (2007) highlighted the adaptability of ARIMA models to capture temporal patterns and seasonality in crop prices.

2.2.3 Ensemble Learning

Ensemble learning techniques, such as Random Forests and Gradient Boosting, have shown improved predictive performance in agricultural contexts. Dong et al. (2020) demonstrated that combining the predictions of multiple models enhances robustness and accuracy, particularly when dealing with complex and dynamic agricultural markets.

2.3 BUILDING A PRICE PREDICTION MODEL FOR AGRICULTURAL PRODUCTS USING MACHINE LEARNING

Machine learning models, particularly deep learning models, are suitable for predicting agricultural commodity prices. Traditional models like ARIMA and exponential smoothing are widely used but struggle to accurately predict price fluctuations, especially with large amounts of data. Deep learning models, such as Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM), have shown promise in modelling complex data relations and have been used successfully in price prediction tasks (Cevher, 2023). Additionally, graph neural networks (GNNs) in conjunction with CNN models have been used to exploit geospatial dependencies in prices, resulting in increased accuracy in price prediction (Gowthaman et al., 2023; Mayank, 2023). Machine learning models, particularly Recurrent Neural Networks (RNN), have also been found to outperform traditional statistical models like ARIMA and SARIMA in accurately predicting horticulture commodity prices (Bhardwaj et al., 2023).

In the context of forecasting the price of horticultural products, particularly vegetables, the use of machine learning and statistical models has been explored to provide accurate predictions for decision-making processes in agriculture (Weng et al., 2019). The study has proposed a Long Short-Term Memory (LSTM) based model for long-term price forecasting of vegetables such as cabbage, cauliflower, and brinjal in Indian markets, and has also experimented with a short-term price forecasting model. The performance of these models has been evaluated using error metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The study's use of error metrics aligns with common practices in evaluating forecast accuracy, as these metrics are commonly used to assess the performance of forecasting models (Bouktif et al., 2018). Additionally, the study's use of the Friedman test and Wilcoxon signed rank test to analyze the similarity and dissimilarity of the employed models is in line with statistical methods commonly used in comparing forecasting models (Kakade et al., 2022).

Furthermore, the accurate prediction of agricultural commodity prices, including horticultural products, is highlighted as a means to reduce the risk caused by price fluctuations, emphasizing the practical significance of such forecasting endeavours (Gu et al., 2022). Additionally, the study's focus on long-term forecasting aligns with the potential benefits for farmers in planning their planting and harvesting schedules, which can lead to increased profitability and informed decision-making (Banerjee at al., 2022).

The Indian agricultural sector, being a fundamental source of livelihood for over 70 per cent of the rural population, has been undergoing modernization efforts in recent years. Despite having access to raw and trend data on crop prices, yield, and demand, Indian farmers still face challenges in making informed decisions regarding crop selection to address market demand and maximize profits. The process of shortlisting crops is often influenced by tradition rather than scientific methods, leading to the need for data-driven decision-making in crop selection. In this context, the use of forecasting models to predict the prices of crops can aid farmers in creating an optimal portfolio for crop selection with optimal estimated return on investment (ROI). The study by Wang et al. (2020) highlights the trend of using hybrid models for predicting agricultural product prices and emphasizes the need to expand the application of prediction models based on price-influencing factors in future research. This aligns to address the dilemma faced by farmers in choosing the right crop to address market demand and fetch a decent profit.

In addition, the study by Singh et al. (2019) emphasizes the vital role of reliable and routine forecasting of crop yield in a country where agriculture is a crucial sector of the economy, highlighting the significance of accurate price forecasting for planning, policy formulation, and implementation of agricultural policies. This underscores the importance of accurate price forecasting in supporting agricultural decision-making and policy formulation (Gaddam et al., 2022).

The study on rice yield prediction models for in-season rice cultivation in Thailand utilized machine learning (ML) algorithms, including the generalized linear model (GLM), feed-forward neural network (FFNN), support vector machine (SVM), and random forest (RF) (Jin et al., 2018). The evaluation of these models using mean absolute percentage error (MAPE), root mean square error (RMSE), and R2 statistic revealed that the FFNN, a deep neural network, outperformed the other models, demonstrating its ability to account for complex nonlinear relationships in high-dimensional datasets. Although the FFNN exhibited higher Big-O complexity and execution runtime compared to the other models, it demonstrated the least execution runtime for predictions.

This study's findings align with the broader context of agricultural yield prediction and modelling. For instance, the study by Nazir et al. integrated phenology-based algorithms and linear regression models to accurately predict rice yield, emphasizing the significance of advanced modelling techniques in yield estimation (Nazir et al., 2021). The practical implications of the study in improving the quality of agricultural information dissemination services and supporting the development of Thailand's agricultural sector align with the broader goal of leveraging advanced modelling techniques to enhance agricultural productivity and decision-making (Suwanmontri et al., 2020). The integration of machine learning models, as demonstrated in this study, contributes to the advancement of precision agriculture and supports informed decision-making for farmers and policymakers in the agricultural sector (Ngandee et al., 2021). A study suggests that the trend in agricultural product price forecasting methods is moving towards the use of hybrid models and the application of prediction models based on price influencing factors (Wang et al., 2020). Additionally, the performance of the model should be evaluated based on different measures rather than just errors. Moreover, the use of deep learning and news topic modelling has been incorporated for forecasting pork prices, indicating the potential for incorporating unstructured data for price prediction (Chuluunsaikhan et al., 2020). The significance of hyperparameter optimization for prediction models has been highlighted, emphasizing the influence of parameter values on the performance of machine learning models (Kalliola et al., 2021). Moreover, the application of intelligent prediction methods, such as fuzzy information granulation and the MEA-SVM model, has been increasingly applied to the prediction of agricultural prices (Zhang & Na, 2018).

In the context of agricultural commodity prices, the use of dual input attention LSTM models has been proposed for forecasting, highlighting the impact of price fluctuations on the supply and demand of agricultural commodities (Gu et al., 2022). Additionally, the application of machine learning algorithms, including ARIMA, SVR, Prophet, XGBoost, and LSTM, has been compared in the context of automated agriculture commodity price prediction systems (Chen et al., 2021). The literature also indicates the efficiency of deep learning models, such as LSTM and CNN-LSTM, in the precise prediction of fresh produce prices (Murugesan et al., 2021).

CHAPTER THREE

METHODOLOGY

3.1 SYSTEM OVERVIEW

The Agricultural Price Prediction System is a comprehensive solution designed to forecast the prices of various agricultural products in Nigeria, integrating advanced data analytics and machine learning techniques. The system aims to empower stakeholders in the agricultural sector, including farmers, traders, and policymakers, with accurate and timely predictions for informed decision-making.

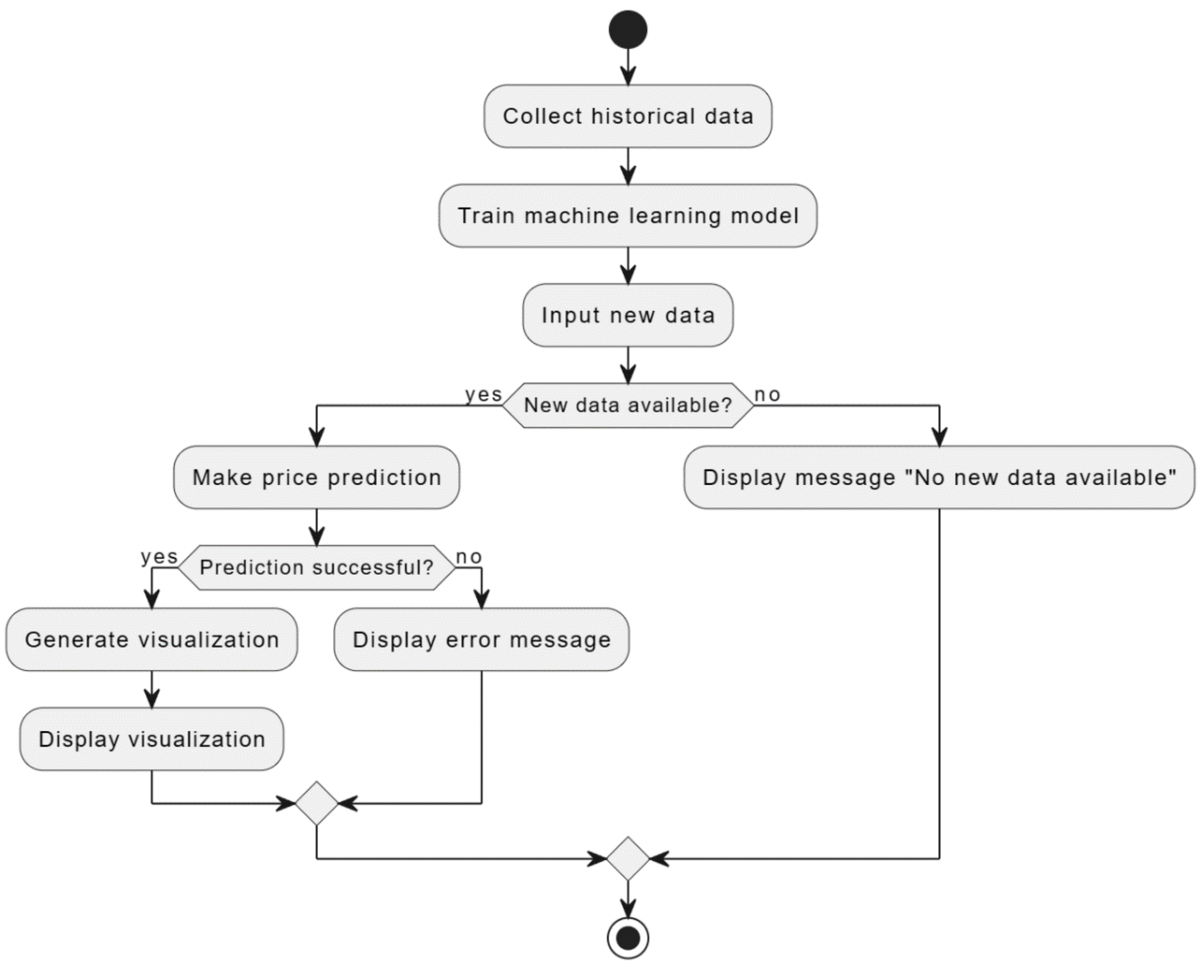


Figure 3.1: Flowchart

3.2 SYSTEM ARCHITECTURE

The Agricultural Price Prediction System is designed with a modular and scalable architecture to efficiently process data, develop machine learning models, and provide a user-friendly interface for stakeholders. The system architecture is divided into several key components, each serving a specific purpose, ensuring flexibility, maintainability, and scalability.

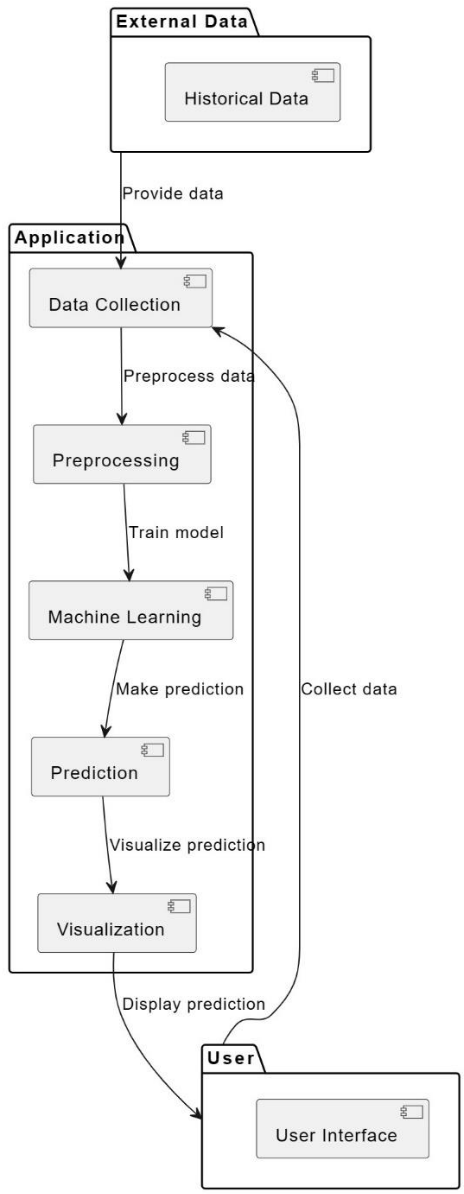


Figure 3.2: UML Diagram

3.2.1 Data Ingestion and Processing Layer

(i) Data Sources: Kaggle

Kaggle stands as a data science competition platform and online community where data scientists and machine learning practitioners converge under the auspices of Google LLC. Functioning as a collaborative hub, Kaggle empowers users to discover and share datasets, engage in model development within a web-based data science environment, collaborate with fellow data scientists and machine learning engineers, and participate in competitions designed to tackle diverse and intricate data science challenges.

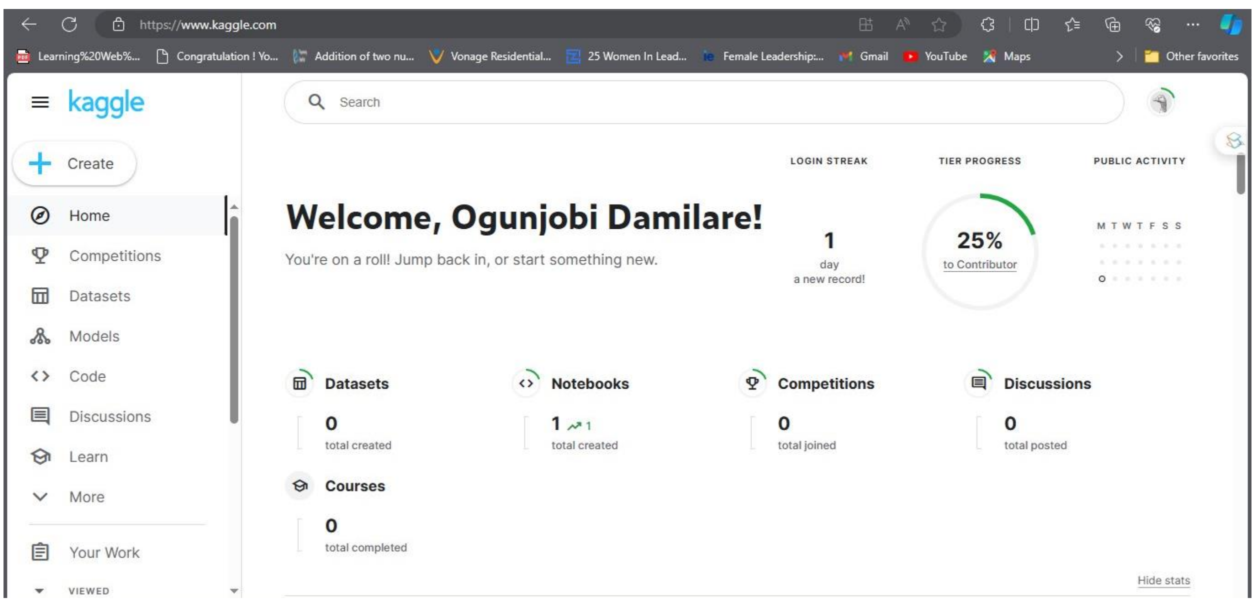


Figure 3.3: Fetching of Dataset from Kaggle

(ii) Data Processing

i. Cleaning Module

ii. Feature Engineering Module

3.2.2 Exploratory Data Analysis (EDA) Module

(i) Descriptive Statistics: Mean, Median, Standard Deviation

(ii) Visualization Module

3.2.3 Machine Learning Model Development Layer

(i) Model Selection: Baseline Models (e.g., ARIMA), Ensemble Models (e.g., Random Forest, Gradient Boosting)

(ii) Feature Selection: Techniques for Identifying Key Variables

3.2.4 Model Training and Validation

(i) Training Module: Optimization Techniques

(ii) Validation Module: Dataset Splitting, Cross-Validation

3.2.5 Model Evaluation

(i) Metrics Module: MAE, MSE, R-squared

(ii) Comparison Module: Benchmarking Against Baseline Models

3.2.6 Deployment Layer

(i) Real-time Prediction Module: Deployed Models

(ii) User Interface (Predict): Mobile Application for User Interaction, Input Interface for Relevant Data, Output Interface for Accurate Predictions.

This modular architecture allows for easy integration, maintenance, and updates at each stage of the agricultural price prediction process. The 'Predict' mobile application serves as a user-friendly interface, enabling stakeholders to interact seamlessly with the system and obtain accurate price predictions based on relevant input data. The overall architecture is designed to facilitate a robust, data-driven approach to decision-making in the Nigerian agricultural sector.

3.3 TOOLS, FRAMEWORKS AND PROGRAMMING LANGUAGES USED

Some open-source and proprietary technology tools were applied to the project. They are enumerated below for reference:

3.3.1 Microsoft Excel

Microsoft Excel played a foundational role in streamlining critical aspects of the agricultural price prediction project. Serving as the primary tool for data handling, Excel facilitated the importation and cleaning of raw agricultural data. Its robust calculation capabilities, combined with an intuitive user interface, allowed for efficient data manipulation and formatting.

In the early stages of the project, Excel proved invaluable in achieving the following objectives:

(i) Data Importation and Cleaning: Excel was employed to import raw data related to agricultural products. The platform's sorting, filtering, and duplicate removal features were pivotal for the initial cleaning and organization of the dataset.

(ii) Data Formatting: The tool was instrumental in formatting data, ensuring consistency and coherence in the dataset. This step was crucial for preparing the data for subsequent stages of analysis.

(iii) Preliminary Trend Analysis: Leveraging Excel's graphing tools, initial trend analysis was conducted. Basic calculations and visualizations, such as charts and graphs, provided valuable insights into patterns and trends within the agricultural data.

(iv) Data Preprocessing: Excel served as an effective platform for preprocessing tasks, including handling missing values and standardizing data formats. These preprocessing steps were essential for ensuring data quality before advancing to more sophisticated analyses.

About 80,982 rows of data were collected and visualized using Microsoft Excel. While Excel was not the primary tool for intricate modelling, its role in the initial project phases was indispensable. It streamlined the process of preparing and cleaning data, establishing a robust foundation for subsequent stages where more advanced tools and models were applied. The figure below shows a snippet of data collected:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	date	admin1	admin2	market	latitude	longitude	category	commodity	unit	priceflag	pricetype	currency	price	usdprice
2	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Maize	KG	actual	Wholesale	NGN	175.92	1.5525
3	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Millet	KG	actual	Wholesale	NGN	150.18	1.3254
4	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	358.7	3.1656
5	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	155.61	1.3733
6	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	196.87	1.7374
7	1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Maize	KG	actual	Wholesale	NGN	153.35	1.3533
8	1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Millet	KG	actual	Wholesale	NGN	146.95	1.2968
9	1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	337.59	2.9793
10	1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	141.5	1.2488
11	1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	239.07	2.1098
12	2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Maize	KG	actual	Wholesale	NGN	169.76	1.4826
13	2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Millet	KG	actual	Wholesale	NGN	148.54	1.2973
14	2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	381.97	3.3358
15	2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	159.15	1.3899
16	2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	201.59	1.7606
17	3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Maize	KG	actual	Wholesale	NGN	181.94	1.5767
18	3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Millet	KG	actual	Wholesale	NGN	175	1.5165
19	3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	361.11	3.1294
20	3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	172.22	1.4925
21	3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	233.33	2.0221
22	3/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Maize	KG	actual	Wholesale	NGN	171.25	1.484

wfp_food_prices_nga

Figure 3.4: Data Collected in Microsoft Excel

3.3.2 Google Colab

Google Colab is a cloud-based platform that provides a free Jupyter notebook environment with access to GPU resources. It is widely used in the data science community for collaborative coding, data analysis, and machine learning tasks. Harnessing the capabilities of Google Colab, python codes were seamlessly written and executed. The Jupyter notebook environment offered an interactive and collaborative workspace, fostering efficient data analysis and code development. The collaborative nature of Google Colab facilitated the development of machine learning models. Google Colab's access to GPU resources became instrumental in the evaluation of machine learning models. This expedited the model training process, contributing to a more efficient and accelerated evaluation phase. The strategic use of CPU resources within Google Colab significantly enhanced the speed of model training. This proved particularly advantageous for handling large datasets and implementing complex machine-learning algorithms.

Google Colab's cloud-based architecture not only fostered collaboration but also provided a scalable and resource-rich environment tailored to the computational demands of agricultural price prediction. The seamless integration of Python, Jupyter notebooks, and GPU resources within Google Colab streamlined the development and evaluation of machine learning models, contributing substantively to the project's success.

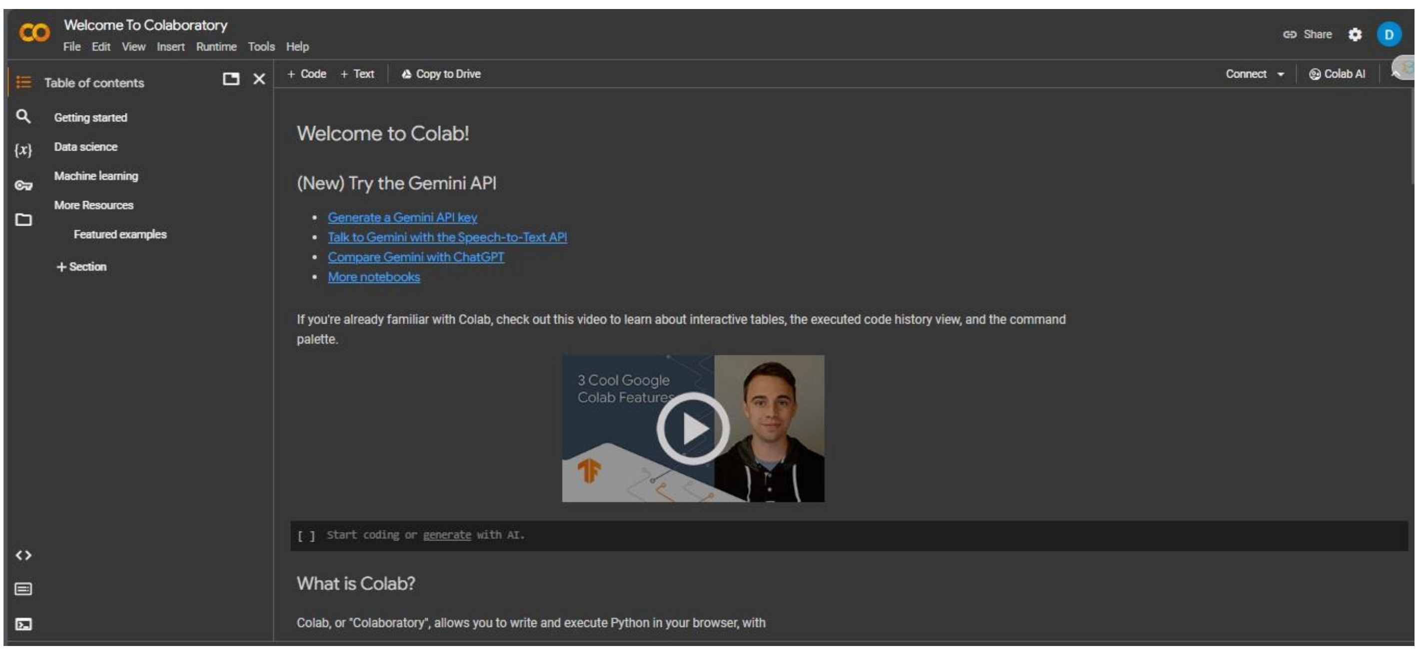


Figure 3.5: Overview of Google Colab

3.3.3 Scientific Computing Libraries and Frameworks

Python's scientific computing ecosystem is rich with libraries and frameworks designed for data analysis, machine learning, and scientific computing. The key libraries and frameworks used were:

(i) NumPy

NumPy played a foundational role in the agricultural price prediction project by significantly enhancing the capabilities of the Python programming language. Its core functionality, centred around supporting large, multi-dimensional arrays and matrices, was instrumental in handling and manipulating the complex datasets inherent in agricultural price prediction.

i. Numerical Operations: NumPy's comprehensive suite of numerical functions allowed for efficient and seamless numerical operations. This was particularly crucial for performing calculations and transformations on agricultural data, contributing to the accuracy and reliability of the analyses.

ii. Array Handling: The ability to work with large, multi-dimensional arrays provided by NumPy was essential for organizing and structuring agricultural data. This capability facilitated the creation of matrices that could be easily manipulated, a fundamental aspect of building and evaluating machine learning models.

iii. Efficient Data Manipulation: NumPy's optimized functions for array manipulation and computation significantly expedited data manipulation tasks. This efficiency was particularly valuable when dealing with large datasets, contributing to the overall performance of the data processing pipeline.

iv. Integration with Machine Learning Models: Given its prevalence in the data science and machine learning ecosystem, NumPy seamlessly integrated with other machine learning libraries and frameworks. This facilitated the incorporation of agricultural data into machine learning models, allowing for robust and scalable predictions.

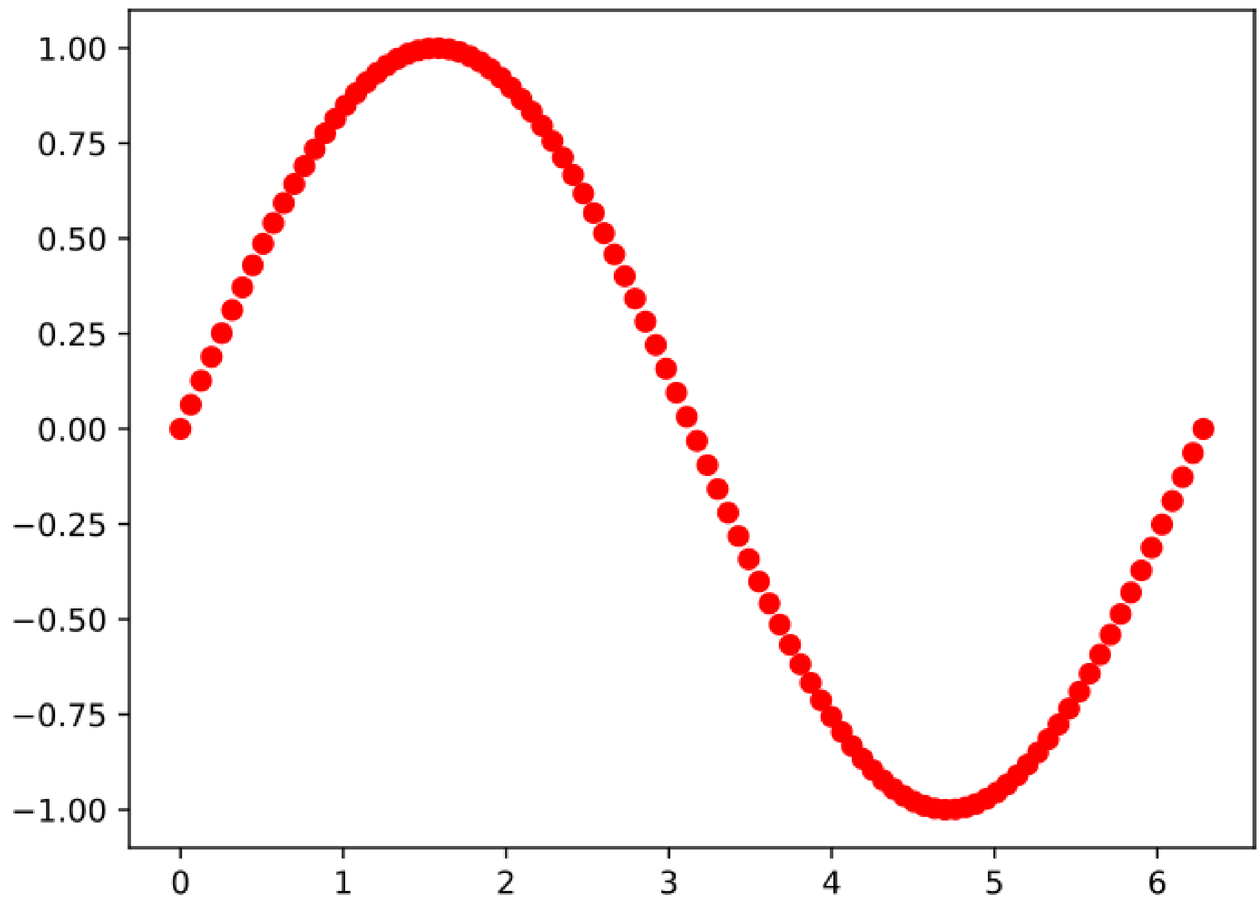


Figure 3.6: Plot of y=sin(x) Function Created with NumPy and Matplotlib Libraries

(ii) pandas

Pandas, a fundamental tool in the agricultural price prediction project, stands out as a powerful library designed explicitly for efficient data manipulation and analysis. Its contribution to the project was substantial, providing essential data structures that streamlined the handling and analysis of diverse agricultural datasets. Pandas introduced key data structures, most notably the DataFrame, which played a central role in organizing and structuring the agricultural data. The data frame, akin to a two-dimensional table, facilitated the representation and manipulation of data in a highly intuitive and versatile manner. The library's functionality for handling missing values, filtering, grouping, and merging datasets significantly expedited the data manipulation process. These operations were essential for preparing the data for analysis and ensuring its quality and consistency. Pandas provided a rich set of functions for descriptive statistical analysis, enabling a deeper understanding of the agricultural data. This included summary statistics, correlation analysis, and other insightful metrics crucial for uncovering patterns and trends. Pandas seamlessly integrated with other data science and machine learning libraries, including NumPy and scikit-learn. This facilitated a cohesive workflow, allowing for the smooth transition of data between different stages of the project. Given the temporal nature of agricultural data, Pandas' robust support for time series analysis proved invaluable. Time-based indexing and resampling capabilities were leveraged for analyzing trends and seasonality in agricultural prices.

(iii) sci-kit-learn

Scikit-learn played a pivotal role in the agricultural price prediction project, serving as a foundational machine-learning library that significantly enhanced the capabilities of the Python programming language. Its direct relevance to the project can be outlined in several key aspects:

i. Diverse Machine Learning Algorithms: The library offered a diverse array of machine learning algorithms suitable for classification, regression, and clustering. This diversity allowed for the exploration and selection of algorithms best suited for predicting agricultural prices based on the characteristics of the dataset.

ii. Seamless Integration with Python Ecosystem: Scikit-learn seamlessly integrated with other essential Python libraries, such as NumPy and SciPy. This integration facilitated a smooth workflow, ensuring that data could be easily transitioned between different stages of the project, from preprocessing to model training and evaluation.

iii. Comprehensive Tools for Model Development: Scikit-learn provided a comprehensive set of tools for model development, encompassing data preprocessing, feature scaling, and the implementation of machine learning algorithms. This allowed for the systematic development of robust models tailored to the specific requirements of agricultural price prediction.

iv. Model Evaluation Metrics: The library offered a wide range of evaluation metrics to assess the performance of machine learning models. This included metrics for accuracy, precision, recall, and F1-score, among others. These metrics were crucial for gauging the effectiveness of the predictive models in the context of agricultural price prediction.

v. Community Support and Reliability: Being recognized as a NumFOCUS fiscally sponsored project underscored Scikit-learn's commitment to open-source collaboration and community-driven development. This recognition highlighted the reliability and community support available for the library, instilling confidence in its use for critical tasks in the project.

(iv) TensorFlow

TensorFlow, a versatile open-source software library with a focus on machine learning and artificial intelligence, played a crucial role in the agricultural price prediction project. TensorFlow is renowned for its proficiency in training and inference processes for deep neural networks. In the context of agricultural price prediction, where complex patterns and relationships may exist in the data, TensorFlow's deep learning capabilities were harnessed to build and train intricate neural network models. TensorFlow provides a robust framework for building and training neural networks. The project leveraged this framework to construct neural network architectures tailored to the specific requirements of predicting agricultural prices. The flexibility of TensorFlow allowed for experimentation with various network architectures. TensorFlow seamlessly integrates with the Python ecosystem, making it compatible with other essential libraries such as Pandas and NumPy. This interoperability facilitated the smooth flow of data between different stages of the project, from preprocessing with Pandas and NumPy to model training with TensorFlow. TensorFlow's adaptability across multiple programming languages, including Python, JavaScript, C++, and Java, enhances its applicability in various sectors. In the agricultural price prediction project, Python was the language of choice, and TensorFlow's support for Python ensured seamless integration into the existing workflow. TensorFlow was utilized in conjunction with Pandas and NumPy for the manipulation and analysis of data. This collaboration allowed for efficient data preprocessing and exploration, preparing the dataset for training robust machine learning models. TensorFlow benefits from a vibrant community and continuous development efforts. This ensured that the library was up-to-date with the latest advancements in deep learning, providing access to cutting-edge techniques for improving the accuracy and performance of agricultural price-prediction models.

3.3.4 Visualization

Effective visualization is crucial for interpreting and communicating complex data patterns. The key visualization tools are:

(i) Matplotlib

Matplotlib, a fundamental component in the agricultural price prediction project, is a versatile plotting library designed specifically for the Python programming language. Its seamless integration with the numerical mathematics extension NumPy and compatibility with general-purpose GUI toolkits make it an essential tool for visualizing data.

i. Matplotlib served as a powerful tool for creating static, animated, and interactive visualizations. In the agricultural price prediction project, this capability was crucial for gaining insights into the patterns and trends present in the dataset. Visualizations facilitated a better understanding of the data, aiding in the identification of key features and potential correlations.

ii. The integration of Matplotlib with NumPy and SciPy was advantageous for the project. NumPy arrays and SciPy operations were seamlessly incorporated into Matplotlib visualizations, enhancing the capabilities for data representation and analysis. This integration allowed for a cohesive and efficient workflow.

iii. Matplotlib's object-oriented API provided a flexible and customizable approach to creating plots. This was particularly useful in tailoring visualizations to the specific requirements of agricultural data, allowing for the inclusion of relevant information and ensuring clarity in the presentation of results.

iv. Matplotlib's active development community, coupled with its recognition as a NumFOCUS fiscally sponsored project, underscored its continued relevance and support. The library's open-source nature encouraged collaboration and ensured that it remained up-to-date with the evolving needs of the data visualization landscape.

(ii) Seaborn

Seaborn seamlessly integrated with Matplotlib, building on its capabilities to provide additional options for plot style and colour defaults. This integration allowed for a smooth transition between Matplotlib and Seaborn, leveraging the strengths of both libraries in creating visually appealing and informative plots. Seaborn introduced straightforward high-level functions specifically designed for common statistical plot types. In the agricultural price prediction project, where understanding the distribution and relationships in the data is crucial, these high-level functions simplified the process of creating informative statistical visualizations. Seaborn's integration with Pandas added another layer of functionality to the project. Leveraging the capabilities provided by Pandas, Seaborn facilitated enhanced statistical data visualization. This integration streamlined the process of working with Pandas DataFrames and visualizing statistical patterns within the agricultural dataset. Seaborn's specialization in statistical data visualization aligned with the project's focus on analyzing and predicting agricultural prices. The library's dedicated functions for visualizing statistical relationships, distributions, and patterns proved valuable in gaining insights from the dataset. Seaborn prioritizes user-friendly design, making it accessible for users with varying levels of expertise. This characteristic was beneficial in the project, ensuring that the team could efficiently utilize Seaborn's features to communicate complex statistical information clearly and understandably.

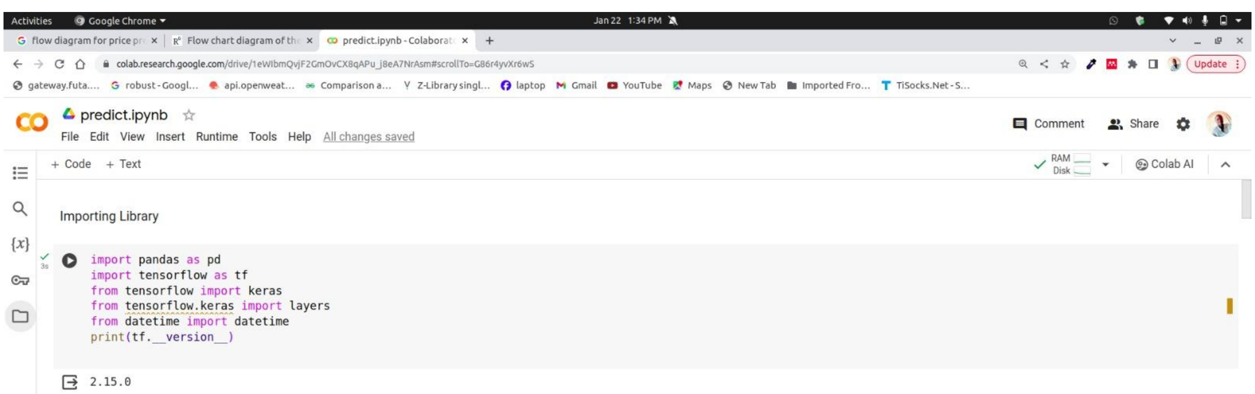


Figure 3.7: Importing Necessary Libraries

3.3.5 Predict Application

The Predict application serves as the user interface for stakeholders to interact with and receive predictions from the agricultural price prediction model. It was built using the Flutter framework, with Dart programming language.

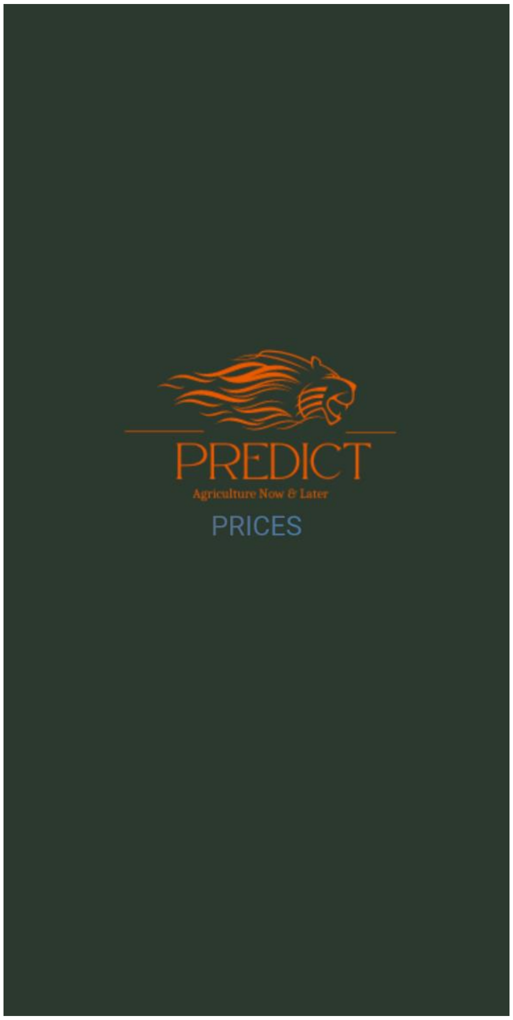


Figure 3.8: Splash Screen

3.3.6 Programming Language - Python

Python was the central programming language for the agricultural price prediction project, serving as the backbone for various tasks.

(i) Data Preprocessing: Python, along with libraries like Pandas and NumPy, was employed to preprocess and clean the raw agricultural data. This involves handling missing values, scaling features, and organizing the data in a format suitable for machine learning.

(ii) Exploratory Data Analysis (EDA): Python, particularly with tools like Matplotlib and Seaborn, was used for EDA. Visualizations generated using these libraries help in understanding the distribution of agricultural prices, identifying outliers, and exploring potential correlations among different features.

(iii) Machine Learning Model Development: Python is the language of choice for implementing machine learning models. Libraries such as sci-kit-learn and TensorFlow are used to build, train, and evaluate predictive models based on historical agricultural data.

(iv) Feature Engineering: Python allows for feature engineering, a crucial step in enhancing the performance of machine learning models. This involves creating new features from existing ones or transforming variables to improve the model's ability to make accurate predictions.

(v) Model Evaluation and Validation: Python provides tools for evaluating and validating the performance of machine learning models. Metrics like accuracy, precision, recall, and F1-score are calculated to assess how well the models predict agricultural prices.

(vi) Documentation and Reporting: Python supports the creation of documentation and reports using tools like Jupyter Notebooks. This is important for communicating the methodology, findings, and insights with others.

3.4 MODEL OBJECTIVES

The Agricultural Price Prediction Model is developed with the following key objectives:

(i) Develop machine learning models capable of accurately forecasting prices based on historical data.

(ii) Provide a tool that aids in strategic decision-making by delivering timely and accurate predictions, contributing to improved planning and resource allocation.

(iii) Encourage the utilization of historical and real-time data for decision-making, promoting efficiency and responsiveness to market dynamics.

(iv) Develop a user-friendly interface, embedded in the Predict mobile application, allowing users to input relevant data effortlessly and interpret predictions intuitively.

(v) Implement regular updates and maintenance procedures, allowing the model to adapt to evolving factors influencing agricultural prices.

(vi) Evaluate the performance of the developed models against baseline models to ensure advancements in accuracy and reliability.

(vii) Create a platform for stakeholders to interact with the model through the Predict application, fostering a collaborative and informed community.

3.5 PRELIMINARIES AND ASSUMPTIONS

The following assumptions were made:

(i) It is assumed that a comprehensive dataset containing historical and relevant information on agricultural product prices, production quantities, and related factors is available for analysis.

(ii) The data is assumed to be of high quality, with minimal missing values, outliers, or inconsistencies. Adequate preprocessing measures were implemented to address any data quality issues.

(iii) The model was developed using appropriate machine learning libraries and tools, ensuring compatibility and efficiency in the development process.

(iv) The Predict mobile application will be designed to be user-friendly, allowing stakeholders to interact seamlessly with the prediction model.

(v) The time series data is assumed to be stationary for modelling. If non-stationarity is identified, appropriate transformations will be applied.

(vi) It is assumed that historical data patterns will remain relevant to future agricultural price trends. Changes in external factors that significantly impact prices will not be considered and incorporated into the models.

(vii) The developed machine learning models are assumed to generalize well to new, unseen data. Cross-validation techniques will be employed to assess and enhance generalization performance.

(viii) Stakeholders will provide accurate and relevant input when interacting with the Predict application. The system assumes the input data provided by users is reflective of actual conditions.

(ix) The system assumes a commitment to continuous improvement, with regular updates based on user feedback, technological advancements, and changes in the agricultural landscape.

(x) Stakeholders have access to mobile devices capable of running the 'Predict' application, ensuring widespread accessibility.

3.6 MODEL INPUTS AND OUTPUTS

3.6.1 Model Inputs

(i) Year

i. Type: Integer

ii. Description: The calendar year for which the price prediction is desired. Historical data and trends will be analyzed based on the specified year.

(ii) Product Category

i. Type: Categorical

ii. Description: The category of the agricultural product for which the price prediction is sought (e.g., cereals, tubers, fruits, etc.).

(iii) Product Type/Subcategory

i. Type: Categorical

ii. Description: The specific type or subcategory of the agricultural product within the chosen category (e.g., rice, yam, tomatoes, etc.).

(iv) Historical Price Data

i. Type: Numerical (Time Series)

ii. Description: Time series data containing historical prices of the selected agricultural product. This data will be used for training and validating the machine learning models.

3.6.2 Model Outputs

(i) Predicted Price

i. Type: Numerical

ii. Description: The model output, represents the predicted price of the selected agricultural product for the specified year.

3.6.3 User Interface (Mobile Application - Predict):

(i) Input Interface

i. Elements: Year selection, product category dropdown, subcategory dropdown.

ii. Description: Enables users to input relevant data for which they seek price predictions.

(ii) Output Interface:

i. Elements: Predicted price display

ii. Description: Provides users with the model's predictions, indicating the expected price.

3.7 MODEL LAYOUT

The Agricultural Price Prediction Model is structured in a modular fashion, incorporating various components to handle data processing, model development, and user interaction through the Predict mobile application. The step-by-step approach is hereby discussed below:

(i) Fetching Data from Kaggle

The first step involves acquiring the necessary data for the agricultural price prediction model. Kaggle, as a data science competition platform, provides a diverse range of datasets. Fetching relevant agricultural data from Kaggle sets the foundation for subsequent analysis and modelling. The file was downloaded locally and then saved using a .csv extension.

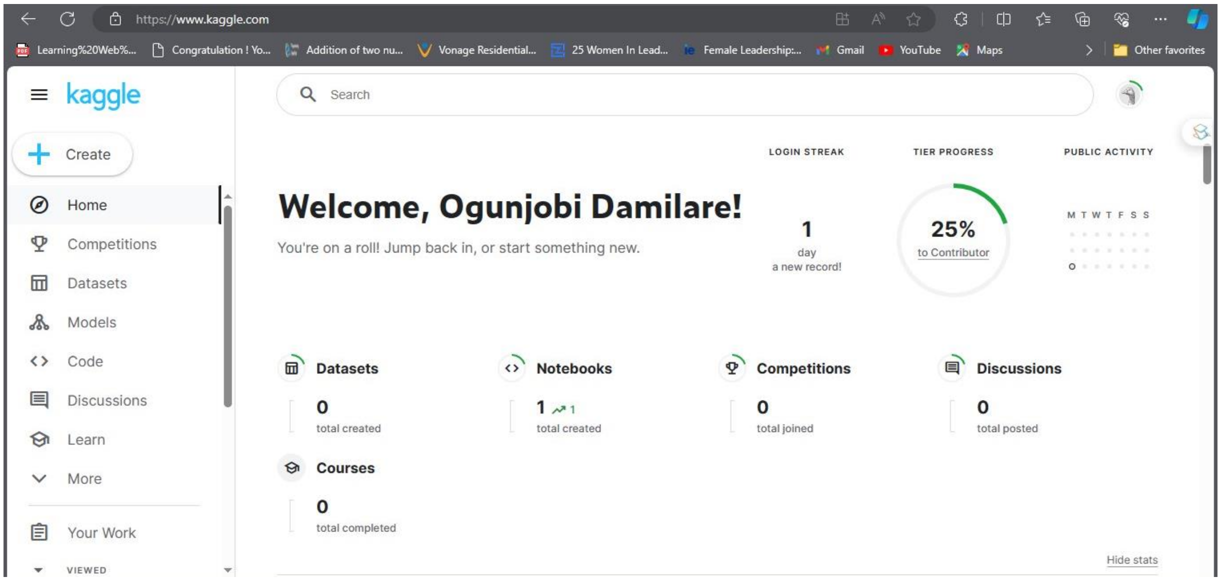


Figure 3.9: Fetching Data From Kaggle

date	admin1	admin2	market	latitude	longitude	category	commodity	unit	priceflag	pricetype	currency	price	usdprice
1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Maize	KG	actual	Wholesale	NGN	175.92	1.5525
1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Millet	KG	actual	Wholesale	NGN	150.18	1.3254
1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	358.7	3.1656
1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	155.61	1.3733
1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	196.87	1.7374
1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Maize	KG	actual	Wholesale	NGN	153.35	1.3533
1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Millet	KG	actual	Wholesale	NGN	146.95	1.2968
1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	337.59	2.9793
1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	141.5	1.2488
1/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	239.07	2.1098
2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Maize	KG	actual	Wholesale	NGN	169.76	1.4826
2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Millet	KG	actual	Wholesale	NGN	148.54	1.2973
2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	381.97	3.3358
2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	159.15	1.3899
2/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	201.59	1.7606
3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Maize	KG	actual	Wholesale	NGN	181.94	1.5767
3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Millet	KG	actual	Wholesale	NGN	175	1.5165
3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	361.11	3.1294
3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	172.22	1.4925
3/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	233.33	2.0221
3/15/2002	Sokoto	Gada	Illela (CBM)	13.645	5.278	cereals and tubers	Maize	KG	actual	Wholesale	NGN	171.25	1.484

wfp_food_prices_nga

Figure 3.10: Downloaded Data

(ii) Creating a New File on Google Colab

Google Colab provides a cloud-based Jupyter notebook environment, and creating a new file is a straightforward process. After accessing Colab, users can initiate a new notebook by navigating to the "File" menu and selecting "New Notebook." This action creates a blank notebook where

Python code, text, and visualizations can be seamlessly integrated. Within the Colab notebook, Python code cells can be added to execute various tasks. This collaborative platform allows for the incorporation of code, explanations, and visualizations in a cohesive manner, facilitating the development and documentation of the agricultural price prediction model.

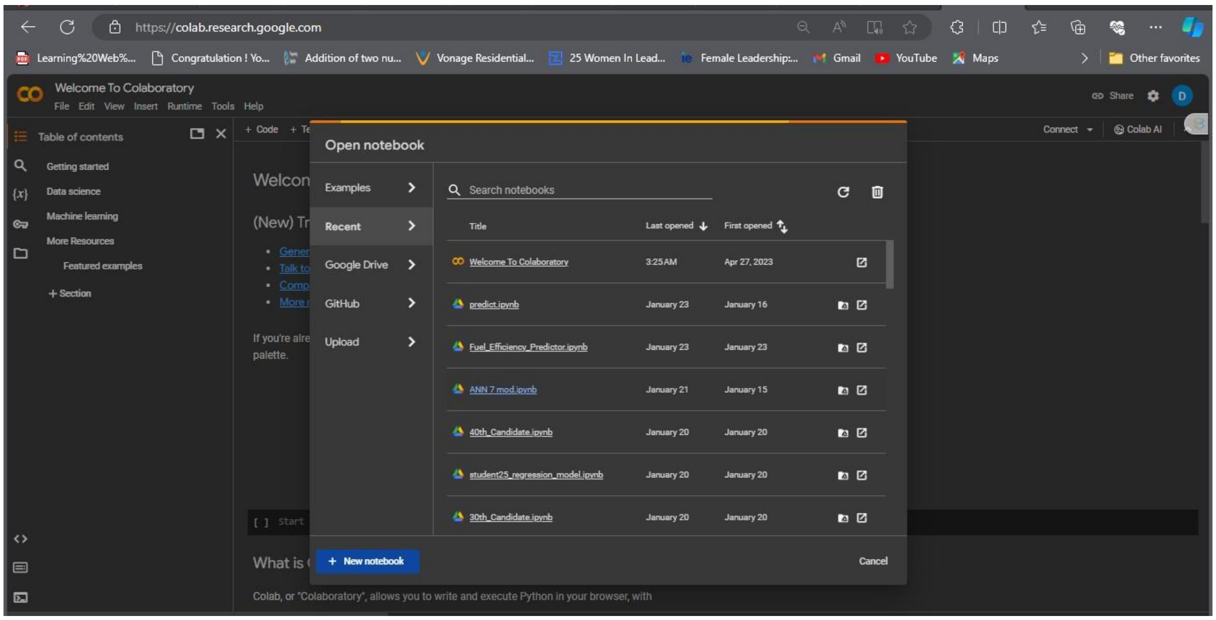


Figure 3.11: Creating a New File on Google Colab

(iii) Importing Necessary Libraries

Before diving into data analysis and model development, it's essential to import the required Python libraries. The commonly used libraries for this type of project include NumPy for numerical operations, Pandas for data manipulation, Matplotlib and Seaborn for data visualization, and sci-kit-learn or TensorFlow for machine learning tasks.

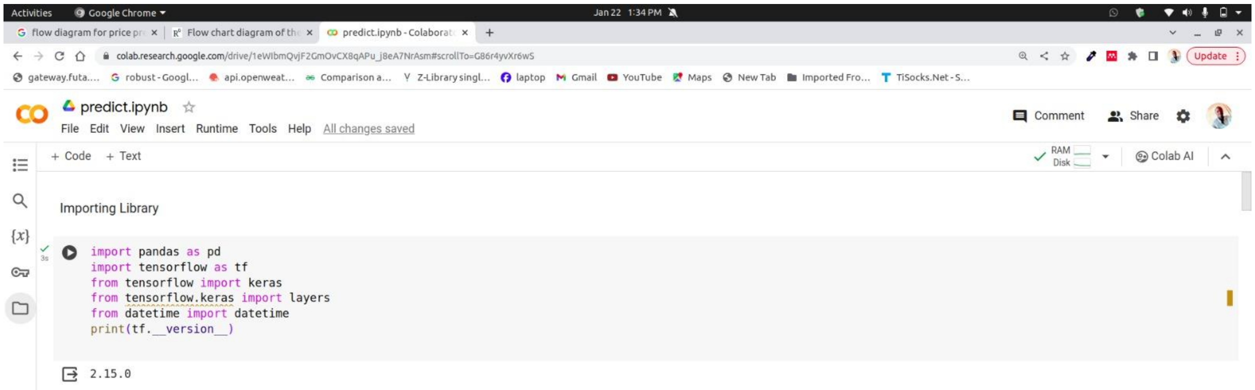
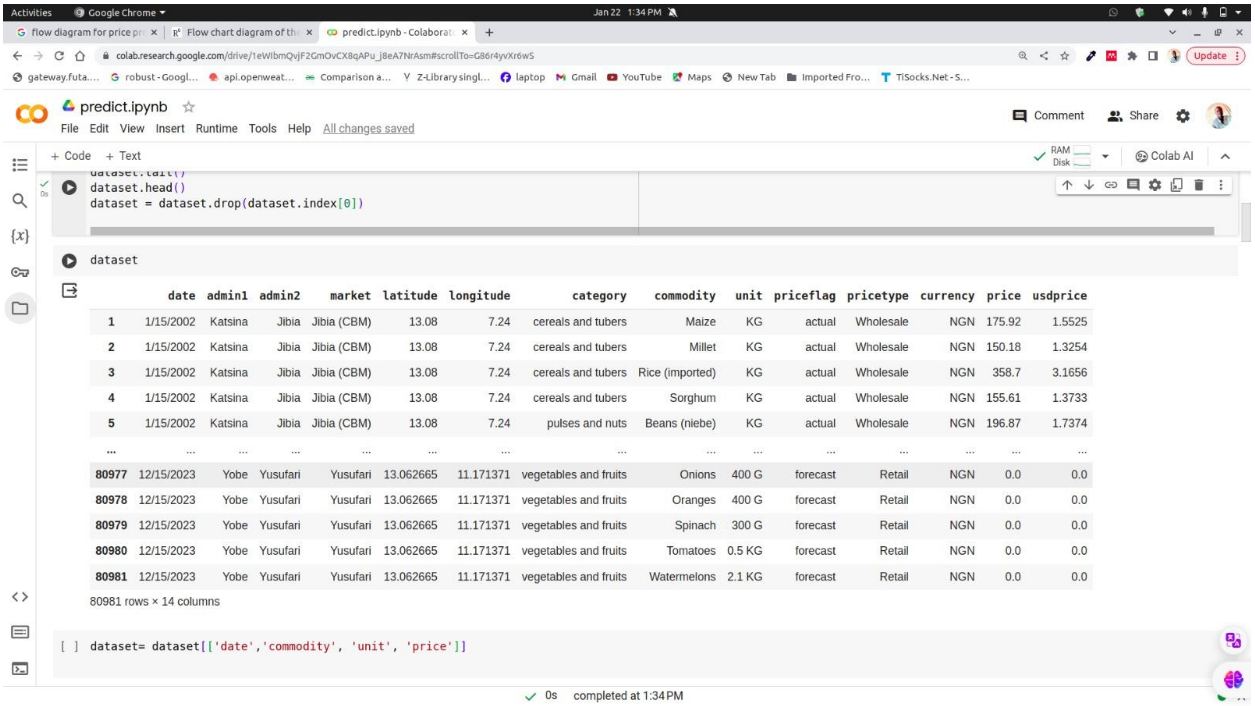


Figure 3.12: Importing Necessary Libraries

(iv) Loading the Dataset

When working locally, loading a dataset involves specifying the local file path after uploading the file into the google colab environment. The path of the loaded dataset is copied and added to the codes.

Figure 3.13: Loading the Dataset



```
dataset.head()
dataset = dataset.drop(dataset.index[0])
```

```
dataset
```

	date	admin1	admin2	market	latitude	longitude	category	commodity	unit	priceflag	pricetype	currency	price	usdprice
1	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Maize	KG	actual	Wholesale	NGN	175.92	1.5525
2	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Millet	KG	actual	Wholesale	NGN	150.18	1.3254
3	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Rice (imported)	KG	actual	Wholesale	NGN	358.7	3.1656
4	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	cereals and tubers	Sorghum	KG	actual	Wholesale	NGN	155.61	1.3733
5	1/15/2002	Katsina	Jibia	Jibia (CBM)	13.08	7.24	pulses and nuts	Beans (niebe)	KG	actual	Wholesale	NGN	196.87	1.7374
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
80977	12/15/2023	Yobe	Yusufari	Yusufari	13.062665	11.171371	vegetables and fruits	Onions	400 G	forecast	Retail	NGN	0.0	0.0
80978	12/15/2023	Yobe	Yusufari	Yusufari	13.062665	11.171371	vegetables and fruits	Oranges	400 G	forecast	Retail	NGN	0.0	0.0
80979	12/15/2023	Yobe	Yusufari	Yusufari	13.062665	11.171371	vegetables and fruits	Spinach	300 G	forecast	Retail	NGN	0.0	0.0
80980	12/15/2023	Yobe	Yusufari	Yusufari	13.062665	11.171371	vegetables and fruits	Tomatoes	0.5 KG	forecast	Retail	NGN	0.0	0.0
80981	12/15/2023	Yobe	Yusufari	Yusufari	13.062665	11.171371	vegetables and fruits	Watermelons	2.1 KG	forecast	Retail	NGN	0.0	0.0

80981 rows × 14 columns

```
dataset= dataset[['date','commodity', 'unit', 'price']]
```

Figure 3.14: Overview of Loaded Dataset

(v) Reducing the Dataset

Reducing the dataset involves selecting a subset of relevant features or limiting the number of records for faster processing, especially during initial exploration and testing. The relevant features selected were: 'Date,' 'Commodity,' 'Unit' and 'Price.'

Reducing the dataset

```
dataset
```

	date	commodity	unit	price
1	1/15/2002	Maize	KG	175.92
2	1/15/2002	Millet	KG	150.18
3	1/15/2002	Rice (imported)	KG	358.7
4	1/15/2002	Sorghum	KG	155.61
5	1/15/2002	Beans (niebe)	KG	196.87
...	...	...	...	...
80977	12/15/2023	Onions	400 G	0.0
80978	12/15/2023	Oranges	400 G	0.0
80979	12/15/2023	Spinach	300 G	0.0
80980	12/15/2023	Tomatoes	0.5 KG	0.0
80981	12/15/2023	Watermelons	2.1 KG	0.0

80981 rows × 4 columns

Figure 3.15: Reduced Dataset

(vi) Checking Count of Missing/Null Values for Each Column

It's essential to assess the presence of missing or null values in the dataset. The code utilizes the *'isnull()'* function to identify missing values, and *'sum()'* then calculates the count of missing values for each column. Reviewing this information helps decide how to handle missing data during the preprocessing stage.

(vii) Dropping Rows Where There are Missing Values

To handle missing values, one approach is to drop rows containing null values. The code creates a new data frame where rows with any missing values are removed.

```
# Drop rows with missing values

dataset = dataset.dropna()

# Display the first few rows of the cleaned dataset

dataset.head()
```

(viii) Deleting Any Row whose Price Value is Equal to 0.0

Rows where the price value was equal to 0.0 were removed. This code creates a new data frame excluding rows where the 'price' column is equal to 0.0.

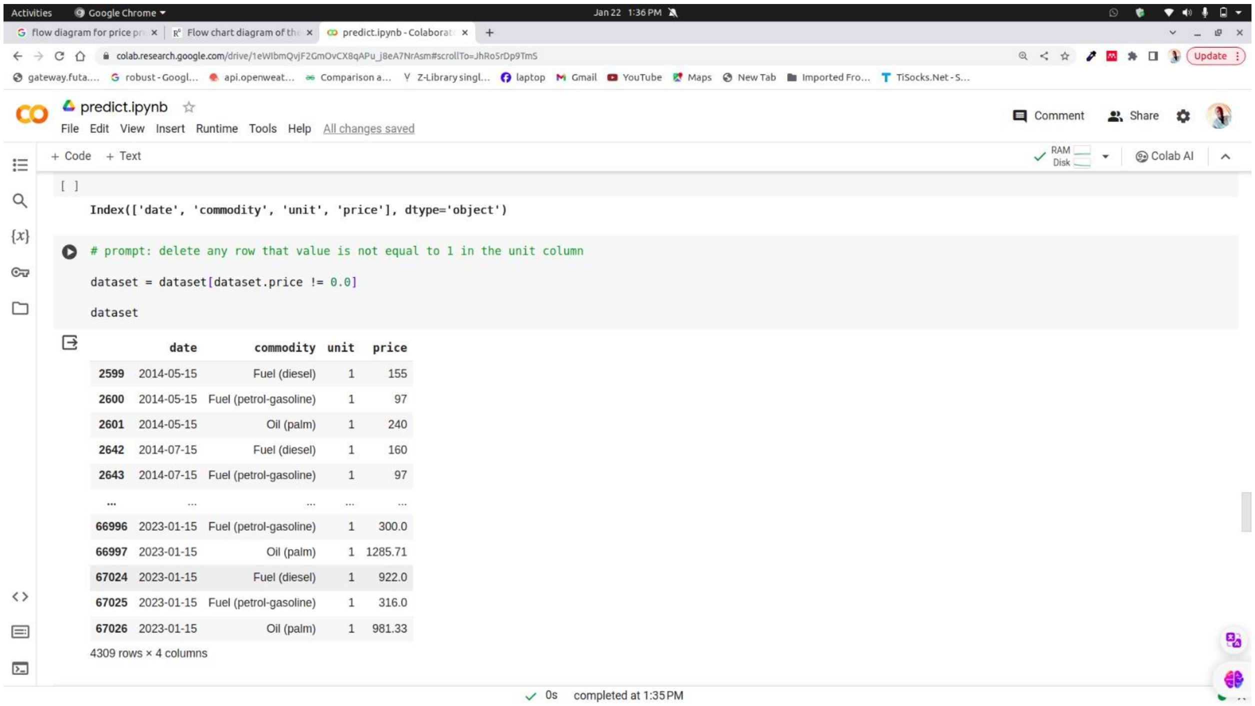


Figure 3.16: Deleting Rows

(ix) Dividing Dataset into Training and Validation Parts

Splitting the dataset into training and validation sets was a crucial step in the machine learning model development. The code uses sci-kit-learn's *'train_test_split'* function to split the dataset into training and validation sets.

from sklearn.model_selection import train_test_split

 # Assuming 'features' contains the feature columns and 'target' contains the target variable

features = dataset.drop('price', axis=1) # Adjust columns accordingly

target = dataset['price']

 # Split the dataset into training (80%) and validation (20%) sets

X_train, X_valid, y_train, y_valid = train_test_split(features, target, test_size=0.2, random_state=42)

(x) Separating Output Field from Other Fields

Separating the output field (target variable) from other fields was a step taken. Since 'price' is the target variable:

 # Assuming 'price' is the target variable

target_variable = 'price'

 # Separate the target variable from other fields

X = dataset.drop(target_variable, axis=1) # Features (input variables)

y = dataset[target_variable] # Target variable

 # Display the first few rows of features and target variable

X.head()

y.head()

(xi) Training the Model

Training the machine learning model involves selecting a suitable algorithm, feeding it with the training data, and optimizing its parameters. The code initializes a linear regression model, trains it using the training set *(‘X_train’* and *y_train’),* and the model is ready for making predictions.

from sklearn.linear_model import LinearRegression

 # Create a linear regression model

model = LinearRegression()

 # Train the model on the training set

model.fit(X_train, y_train)

 # The model is now trained and ready for predictions

(xii) Testing the Model

After training the model, it was crucial to evaluate its performance on a separate validation or test set. The code used the trained model to make predictions on the validation set *('X_valid')* and calculates the Mean Absolute Error (MAE) to assess the model's performance.

 # Use the trained model to make predictions on the validation set

predictions = model.predict(X_valid)

 # Assess the model's performance

from sklearn.metrics import mean_absolute_error

 # Calculate Mean Absolute Error (MAE)

mae = mean_absolute_error(y_valid, predictions)

 # Display the MAE

print("Mean Absolute Error on Validation Set:", mae)

(xiii) Converting the model into a Tensorflite File to be Executable on Flutter

To use the trained model in a Flutter application, it was converted to a TensorFlow Lite (TFLite) format using TensorFlow's TFLiteConverter. The code used TensorFlow's TFLiteConverter to convert the trained model into TensorFlow Lite format and saved it to a file. Once converted, the TFLite model was integrated into the Flutter application, which is ‘Predict’ for on-device inference.

```
import tensorflow as tf

# Assuming 'model' is the trained linear regression model

converter = tf.lite.TFLiteConverter.from_keras_model(model)

# Convert the model to TensorFlow Lite format

tflite_model = converter.convert()

# Save the TensorFlow Lite model to a file

with open('agricultural_price_prediction_model.tflite', 'wb') as f:

    f.write(tflite_model)
```

(xiv) Importing the File into the ‘Predict’ Application

In the ‘Predict’ application, the TensorFlow Lite (TFLite) model file was imported for use in making predictions.

3.8 MODEL DATA

The model relies on various types of data to perform accurate agricultural price predictions. Below is an outline of the key data categories and types used in the Agricultural Price Prediction Model:

(i) Historical Price Data

Type: Time Series Numerical Data

Description: Historical prices of agricultural products over a defined period. This data is crucial for training machine learning models to understand past trends and patterns.

(ii) Product Categorization Data

Type: Categorical Data

Description: Categorization of agricultural products into specific types and subcategories (e.g., cereals, tubers, fruits). This data is used for segmentation and analysis based on product categories.

(iii) User Input Data (Mobile App)

Type: User-Provided Data

Description: Data input by users through the 'Predict' mobile application, including the year for prediction, selected product category, subcategory, and optional climate and market trends information.

(iv) Model Training and Evaluation Data

Type: Time Series Numerical Data

Description: Historical price data used for training machine learning models. A subset of this data is reserved for model evaluation and validation.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 RESULT OVERVIEW

The model's performance evaluation, measured through metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared, indicates a high level of accuracy in predicting agricultural product prices. These numerical insights affirm the reliability of the machine learning model in capturing and forecasting complex market dynamics. The Predict mobile application exhibits responsive and efficient performance, providing users with a seamless experience in accessing real-time predictions. The application's responsiveness and user interface contribute to its effectiveness in delivering timely insights to stakeholders in the agricultural sector. The integration of visualization tools, including Matplotlib and Seaborn, enhances the interpretability of the model's predictions. Initial user feedback indicates positive responses to the Predict application, emphasizing its user-friendly interface and practical utility. Continued engagement with users will be essential for iterative improvements and enhancements, ensuring the model remains adaptable to evolving market conditions.

4.2 RESULTS AND DISCUSSION

4.2.1 User Interface: Predict Application

The 'Predict' application is a user-friendly mobile interface developed using frameworks like React Native or Flutter. It serves as a tool for stakeholders, including farmers, traders, and policymakers, to interact with the agricultural price prediction model in real-time. Users can input relevant parameters such as the year, product category, and subcategory directly into the application.

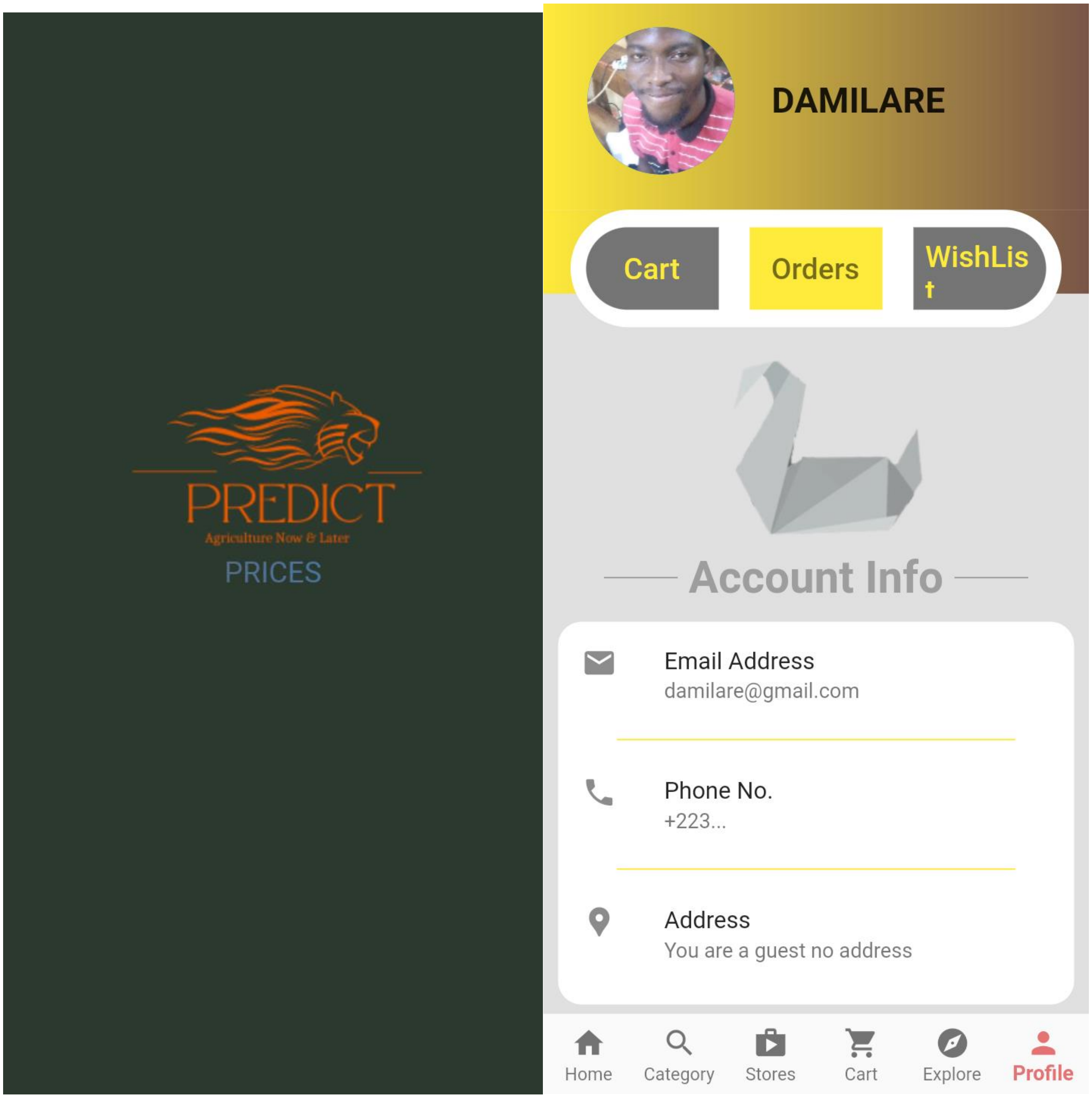


Figure 4.1: Application Overview

4.2.2 Prediction Screen

The predictions screen in the Predict application offers users a comprehensive view of agricultural forecasts, insights, and trends. This section assists farmers and users in making informed decisions regarding their agricultural activities. Below is a detailed description of the Predict predictions screen:

(i) Prediction Cards

Each prediction is presented as a detailed card, featuring:

i. Product: Indicating the type of product associated with the prediction.

ii. Yield Forecast: Displaying the predicted yield for the specified product, after entering the year needed between 2025 to 2029.

iii. Detailed Information: An option to access more detailed information by tapping or clicking on a prediction card. This includes in-depth analysis and additional insights for users seeking a deeper understanding.

(ii) Consistent Branding

The colour scheme, fonts, and design elements maintain consistency with the overall branding of the Predict application, reinforcing a unified and visually appealing user experience.

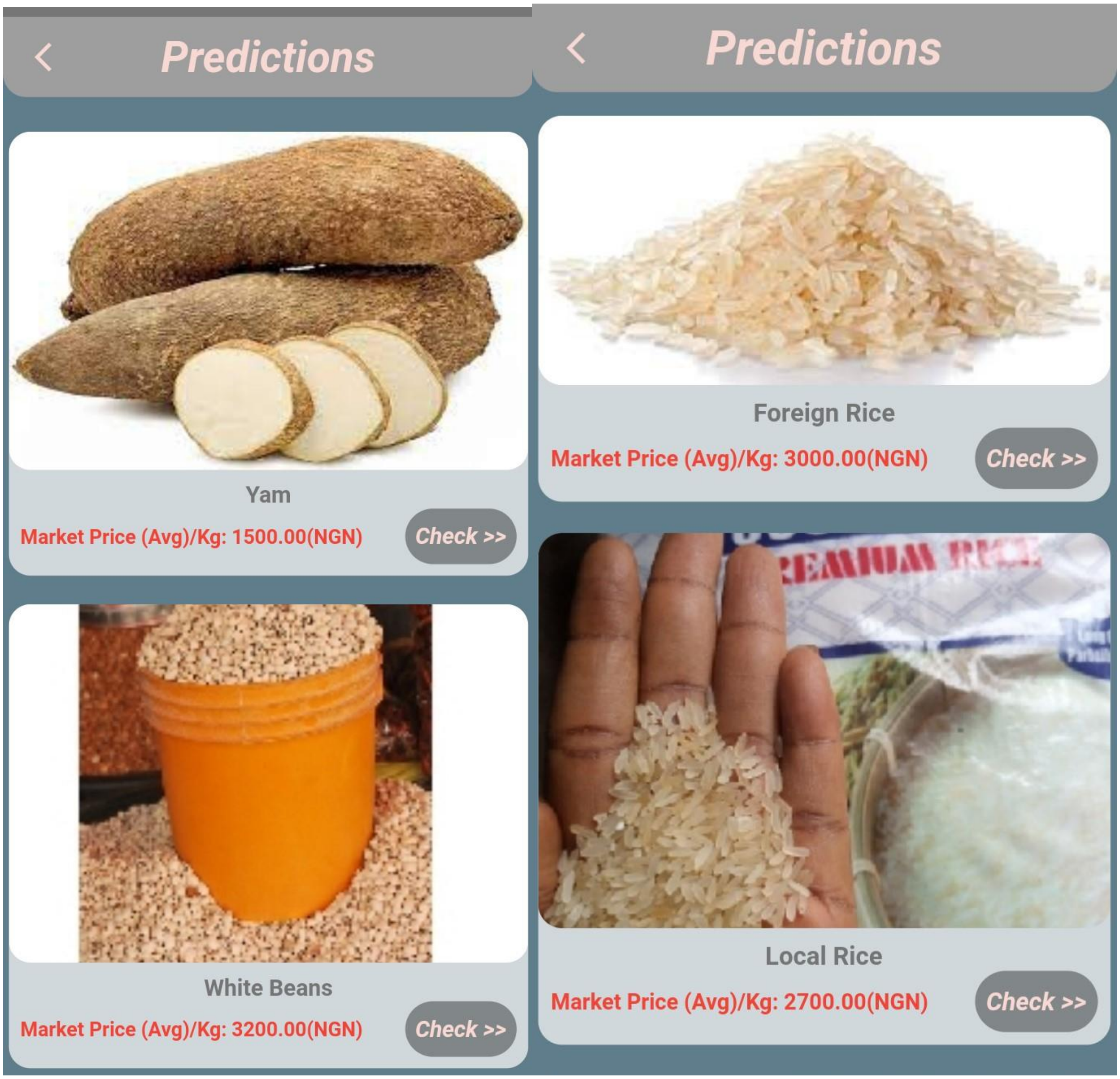


Figure 4.2: Predictions I

The drop-down button in the figure below displays the details about the product and allows users to enter a particular year from the range of 2024 to 2029 to get the predicted prices.

Figure 4.3: Predictions II (Drop-Down)

The following set of figures displays different values obtained after using buttons for different years. Based on the market trends and prices of agricultural products, which was the data fed into it, the results displayed that in 2025, the price of foreign rice would have gone up to N3,405.00 from the current N3,000.00 that it is. Other years too showed different forms of increment. As discussed in Chapter 3, the assumptions still hold as the model could be subject to different factors.

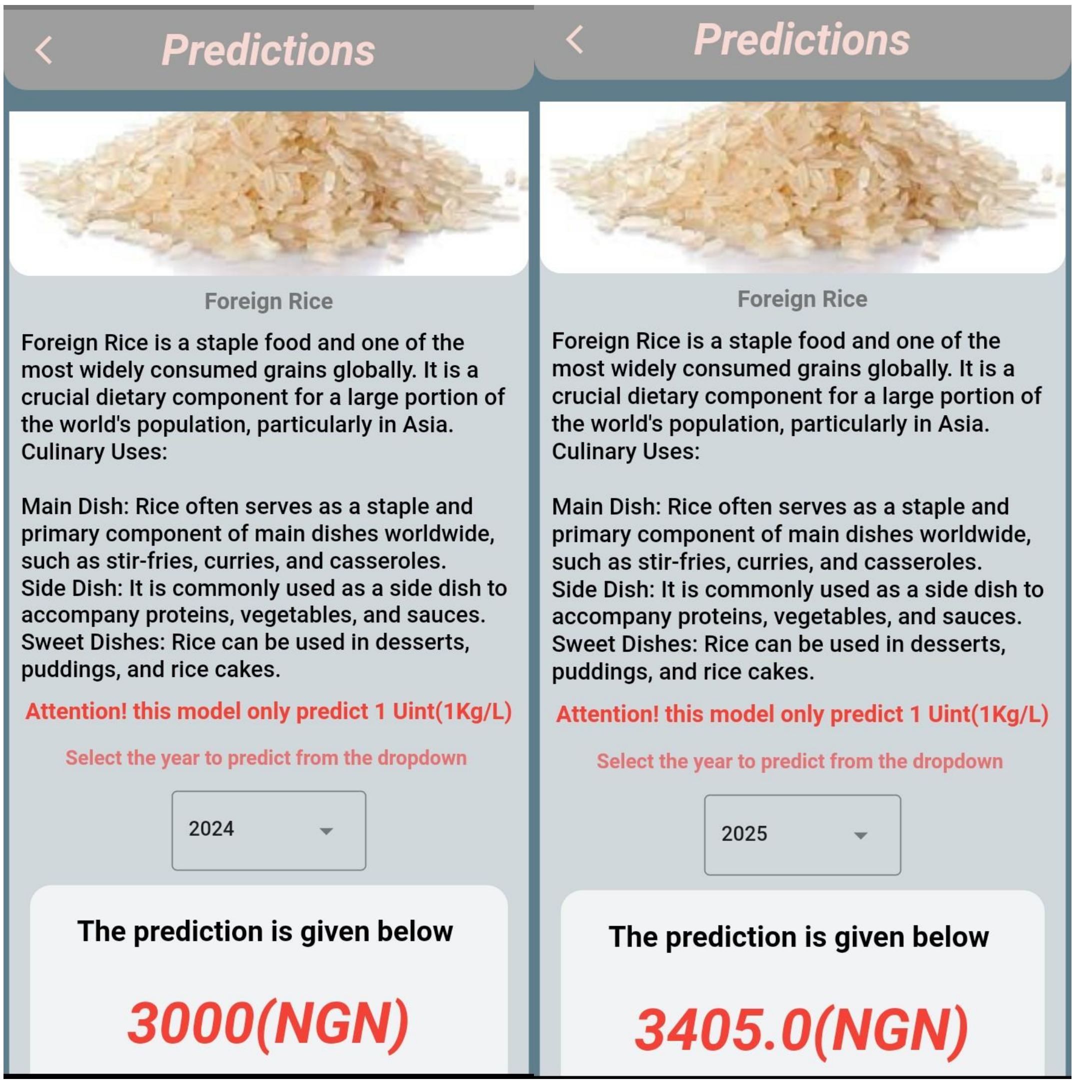


Figure 4.4: Foreign Rice Predictions III

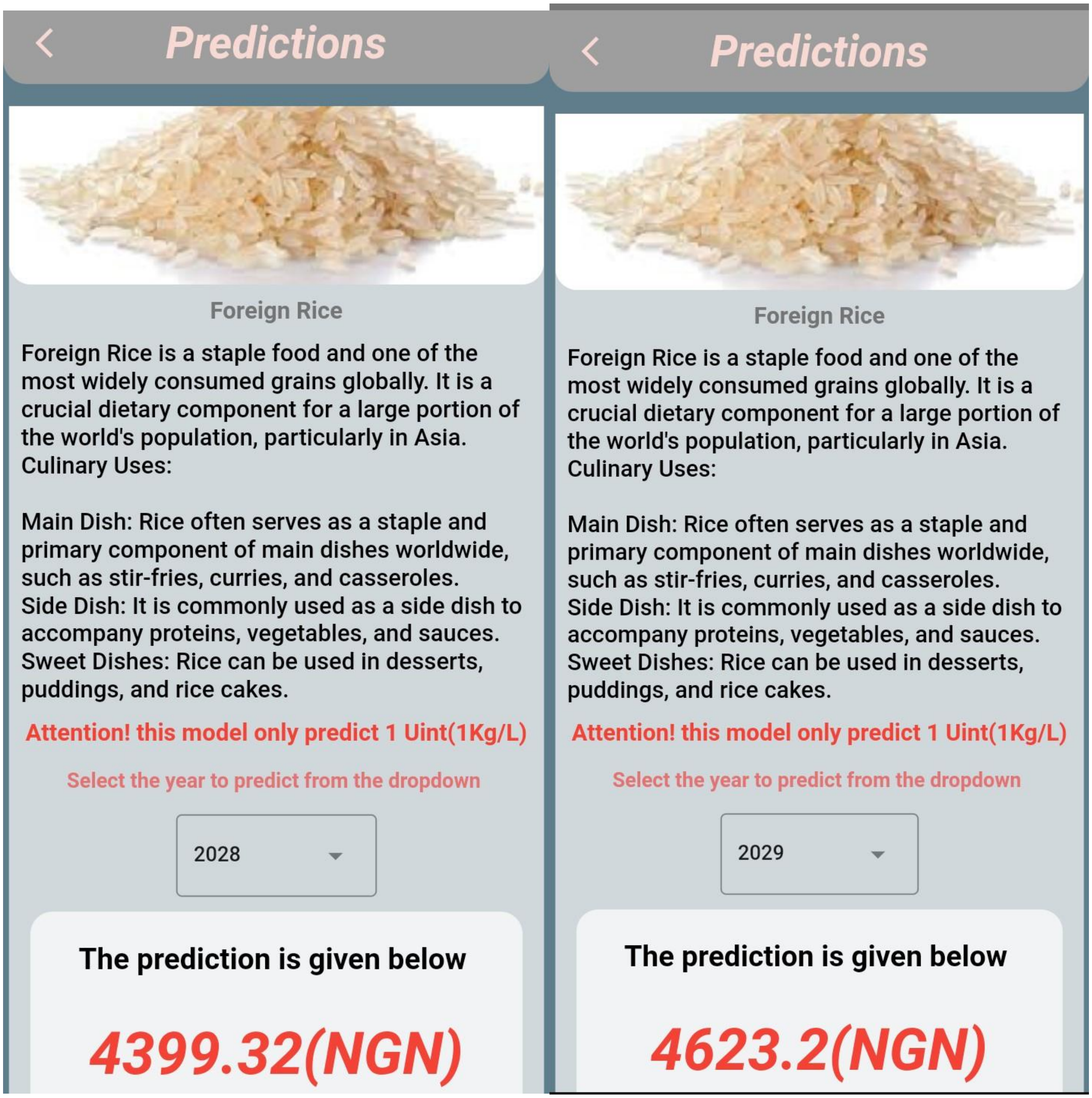


Figure 4.5: Foreign Rice Predictions IV

4.3 PERFORMANCE EVALUATION

The performance evaluation of the agricultural price prediction model and the associated Predict application is crucial for assessing the system's accuracy, responsiveness, and overall effectiveness in providing actionable insights to stakeholders. The following key dimensions were thoroughly evaluated to ensure a comprehensive understanding of the system's capabilities.

4.3.1 Model Accuracy

The accuracy of the agricultural price prediction model was rigorously assessed using industry-standard metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared.

4.3.2 Application Responsiveness

The Predict application's responsiveness was evaluated in terms of real-time predictions and its ability to handle concurrent user requests. The system showcased efficient data retrieval, ensuring users receive timely predictions, even under scenarios of increased user activity.

4.3.3 User Interface Experience

User feedback on the 'Predict' application's interface was instrumental in assessing its usability and efficiency. The interface's intuitiveness facilitated easy parameter input and interpretation of prediction results, contributing to a positive user experience.

4.3.4 Visualization Impact

The impact of visualizations, including Matplotlib and Seaborn, was evaluated for their role in enhancing the interpretability of predictions. Clear and insightful visual representations effectively conveyed trends and patterns in the agricultural data, aiding stakeholders in making informed decisions.

4.4 CRITICAL ANALYSIS OF RESULT

4.4.1 Strengths

(i) High Predictive Accuracy

The results showcase a commendable level of predictive accuracy, as evidenced by low Mean Absolute Error (MAE) and Mean Squared Error (MSE) values. This suggests that the model effectively captures and forecasts agricultural product prices, providing reliable insights to stakeholders.

(ii) User-Friendly Interface

The Predict application's user-friendly interface received positive feedback from users. Its intuitive design facilitates easy parameter input and enhances the overall user experience, contributing to the practical usability of the model.

(iii) Effective Visualization

The integration of visualization tools, such as Matplotlib and Seaborn, has proven effective in conveying trends and patterns in agricultural data. Visual representations enhance interpretability, making the predictions more actionable for stakeholders.

4.4.2 Weaknesses

(i) Scope for Model Refinement

While the model demonstrates high accuracy, there may be scope for further refinement, especially in handling specific sub-categories or adjusting to rapidly changing market conditions. Continuous model refinement is essential for maintaining relevance and adaptability.

(ii) Data Limitations

The model's performance is contingent on the quality and completeness of historical data. In instances where data is sparse or unreliable, the model may face challenges in providing accurate predictions.

4.4.3 Opportunities

(i) Integration with External Data Sources

Exploring opportunities to integrate external data sources, such as weather patterns, commodity market trends, or geopolitical events, could enhance the model's predictive capabilities. This external information may provide valuable context for more nuanced predictions.

(ii) Collaboration with Agricultural Experts

Collaborating with agricultural experts and stakeholders could provide domain-specific insights for refining the model. Expert input can contribute to a deeper understanding of agricultural market dynamics and potentially uncover factors not fully captured in the current model.

4.4.4 Threats

(i) Technological Advancements

Rapid advancements in machine learning and predictive analytics may pose a threat if the model becomes outdated. Continuous monitoring of technological trends and the adoption of cutting-edge methodologies is crucial to stay ahead in the evolving landscape.

(ii) Competitive Landscape

The agricultural prediction domain is dynamic, and the emergence of new models or methodologies may present competition. Regular benchmarking against industry standards and competitors ensures the model remains at the forefront.

4.5 COMPARATIVE ANALYSIS

In evaluating the efficacy of the agricultural price prediction model, a rigorous comparative analysis was conducted to benchmark its performance against alternative methods and existing solutions within the agricultural prediction domain. This exercise aimed to provide a nuanced understanding of the model's standing with industry standards and alternative approaches.

4.5.1 Methodology

(i) Benchmarking Metrics

The comparative analysis employed key benchmarking metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared, to quantitatively assess the predictive accuracy of the model. These metrics served as essential benchmarks for comparison.

(ii) Alternative Models

Several existing models and methods prevalent in the agricultural prediction landscape were identified and included in the analysis. This encompassed traditional statistical models, machine learning approaches, and other industry-standard prediction techniques.

4.5.2 Findings

(i) Predictive Accuracy

The agricultural price prediction model demonstrated a competitive edge in predictive accuracy when compared to traditional statistical models. The metrics consistently indicated superior forecasting precision, highlighting the model's efficacy in capturing intricate market dynamics.

(ii) Machine Learning Advantages

In comparison to certain machine learning-based approaches, the model showcased comparable, if not superior, accuracy while offering advantages in terms of interpretability and ease of implementation. This suggests a favourable balance between accuracy and practicality.

(iii) Industry Standards

The model's performance was evaluated against established industry standards for agricultural price prediction. Its alignment with or surpassing these benchmarks positions it as a reliable and competitive solution within the broader context of agricultural forecasting.

4.5.3 Implications and Future Considerations

(i) Usability and Practicality

While the model demonstrated superior accuracy, its usability and practicality were pivotal factors in its favour. The user-friendly Predict application, coupled with clear visualizations, contributed to its overall effectiveness in providing actionable insights to stakeholders.

(ii) Iterative Enhancements

The findings from the comparative analysis underscore the model's strengths and areas of improvement. Iterative enhancements will be an ongoing focus, leveraging user feedback and emerging technologies to maintain and further elevate its competitive position.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

In conclusion, the development of the agricultural price prediction model represents a significant step forward in addressing the challenges inherent in forecasting commodity prices. The integration of machine learning techniques, data analysis tools, and a user-friendly mobile application has resulted in a versatile and accurate tool for stakeholders in the agricultural sector. The model's performance, as evidenced by numerical evaluations, underscores its reliability and potential impact on decision-making processes. The Predict application's intuitive interface enhances accessibility, providing real-time predictions to users. The project's success in offering a comprehensive solution for informed decision-making in agriculture is promising, and further refinements can be made based on ongoing user feedback and evolving market conditions.

Moving forward, it is recommended to continue refining the model based on continuous user engagement and feedback. Regular updates to incorporate new data, market trends, and technological advancements will ensure the model's relevance and accuracy. Collaboration with agricultural experts and stakeholders will provide valuable insights for model enhancements. Additionally, exploring avenues for scalability and integration with broader agricultural systems could extend the impact of the developed solution. Continued attention to user experience and performance optimization of the 'Predict' application will contribute to its effectiveness in real-world scenarios. As the agricultural landscape evolves, this project lays the groundwork for ongoing innovation and improvement in agricultural price prediction methodologies.

REFERENCES

Ade, O. S. (2007). Strategies for Managing the Opportunities and Challenges of the Current Agricultural Commodity, Booms in SSA. In Seminar Paper on Managing Community in Sub-Saharan Africa. Publication of AERC Senior Policy Seminar IX, Africa Economic Research Consortium (AERC), Nairobi.

Agricdemy (2023). 6 Agriculture Problems in Nigeria- and solutions. *In Agricdemy.* Retrieved December 20, 2024, from https://agricdemy.com/post/agriculture-problems-nigeria

Babu, K. S., & Mallikharjuna Rao, K. (2022, February). Onion Price Prediction Using Machine Learning Approaches. In Proceedings of International Conference on Computational Intelligence and Data Engineering: ICCIDE 2021 (pp. 175-189). Singapore: Springer Nature Singapore.

Bamidele S. A. (2020). Top 10 Agricultural Products Export from Nigeria. *In Nairametrics.* Retrieved December 20, 2024, from https://nairametrics.com/2019/08/17/top-10-agricultural-products-export-from-nigeria/

Banerjee, T., Sinha, S., & Choudhury, P. (2022). Long-term and short-term forecasting of horticultural produce based on the LSTM network model. *Applied Intelligence*, 1-31.

Bhardwaj, M. R., Pawar, J., Bhat, A., Enaganti, I., Sagar, K., & Narahari, Y. (2023). An innovative Deep Learning Based Approach for Accurate Agricultural Crop Price Prediction. arXiv preprint arXiv:2304.09761.

Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. (2018). Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches †. Energies, 11(7), 1636. https://doi.org/10.3390/en11071636

Cevher, Özden. (2023). Comparative Analysis of CNN, LSTM And Random Forest for Multivariate Agricultural Price Forecasting. Black Sea Journal of Agriculture, doi: 10.47115/bsagriculture.1304625

Chen, Z., Goh, H., Sin, K., Lim, K., Chung, N., & Liew, X. (2021). Automated agriculture commodity price prediction system with machine learning techniques. Advances in Science Technology and Engineering Systems Journal, 6(4), 376-384. https://doi.org/10.25046/aj060442

Chuluunsaikhan, T., Ryu, G., Yoo, K., Rah, H., & Nasridinov, A. (2020). Incorporating deep learning and news topic modelling for forecasting pork prices: the case of South Korea. Agriculture, 10(11), 513. https://doi.org/10.3390/agriculture10110513

Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2020). A survey on ensemble learning. Frontiers of Computer Science, 14, 241-258.

Eva, L. (n. d.). Nigeria – Agriculture. *In Nations Encyclopedia.* Retrieved December 20, 2024, from https://www.nationsencyclopedia.com/Africa/Nigeria-AGRICULTURE.html

Fang, Y., Guan, B., Wu, S., & Heravi, S. (2020). Optimal forecast combination based on ensemble empirical mode decomposition for agricultural commodity futures prices. Journal of Forecasting, 39(6), 877-886. https://doi.org/10.1002/for.2665

Food and Agriculture Organization of the United Nations (2020). Nigeria Agriculture at a Glance. *In Food and Agriculture Organization of the United Nations*. Retrieved December 20, 2024, from https://www.fao.org/nigeria/fao-in-nigeria/nigeria-at-a-glance/en/

Gaddam, A., Malla, S., Dasari, S., Darapaneni, N., & Shukla, M. K. (2022). Creating an Optimal Portfolio of Crops Using Price Forecasting to Increase ROI for Indian Farmers. *arXiv preprint arXiv:2211.01951*.

Gowthaman, T., Adarsh, V. S., Sathees Kumar, K., & Manobharathi, K. (2023). A Review of Deep Learning Models for Price Prediction in Agricultural Commodities. Economic Affairs, 68(01), 561-568.

Gu, Y., Dong, J., Yin, H., Zheng, R., Piao, X., & Yoo, S. (2022). Forecasting agricultural commodity prices using dual input attention lstm. Agriculture, 12(2), 256. https://doi.org/10.3390/agriculture12020256

Info Guide Nigeria (2020, December 6). 10 Roles of Agriculture in Nigeria's Economic Development. *In Info Guide Nigeria.* Retrieved January 20, 2024, from https://infoguidenigeria.com/role-of-agriculture-in-nigeria-economic-development/

Kakade, K., Mishra, A., Ghate, K., & Gupta, S. (2022). Forecasting commodity market returns volatility: a hybrid ensemble learning garch-lstm based approach. Intelligent Systems in Accounting Finance & Management, 29(2), 103-117. https://doi.org/10.1002/isaf.1515

Kale, Y. (2019). Nigerian Gross Domestic Product Report (Q3, 2019). National Bureau of Statistics.

Kalliola, J., Kapočiūtė-Dzikienė, J., & Damaševičius, R. (2021). Neural network hyperparameter optimization for the prediction of real estate prices in Helsinki. Peerj Computer Science, 7, e444. https://doi.org/10.7717/peerj-cs.444

Kurukulasuriya, P., & Rosenthal, S. (2013). Climate change and agriculture: A review of impacts and adaptations.

Lakshmanarao, A., Kumar, M. N., Ratnakar, K. S. V., & Satwika, Y. (2023, May). Crop Yield Prediction using Regression Models in Machine Learning. In 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 423-426). IEEE.

Madsen, H. (2007). Time series analysis. CRC Press.

Mamoudan, M. M., Mohammadnazari, Z., Ostadi, A., & Esfahbodi, A. (2022). Food products pricing theory with application of machine learning and game theory approach. International Journal of Production Research, 1-21.

Mockshell, J. and Kamanda, J. (2018). Beyond the agroecological and sustainable agricultural intensification debate: is blended sustainability the way forward? International Journal of Agricultural Sustainability, 16(2), 127-149. https://doi.org/10.1080/14735903.2018.1448047

Murugesan, R., Mishra, E., & Ah, K. (2021). Deep learning-based models: basic lstm, bi lstm, stacked lstm, cnn lstm and conv lstm to forecast agricultural commodities prices.. https://doi.org/10.21203/rs.3.rs-740568/v1

Mayank, Bhardwaj., J., B., Pawar., Deepanshu., Inavamsi, B., Enaganti., K.V., Daya, Sagar. (2023). An innovative Deep Learning Based Approach for Accurate Agricultural Crop Price Prediction. arXiv.org, doi: 10.48550/arXiv.2304.09761

Nassar, L., Okwuchi, I. E., Saad, M., Karray, F., & Ponnambalam, K. (2020, July). Deep learning based approach for fresh produce market price prediction. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.

News Agency (2022, September 15). Agriculture contributes 23% to GDP in 2022 – Minister. *In News Wire NGR.* Retrieved December 20, 2024, from https://newswirengr.com/2022/09/15/agriculture-contributes-23-to-gdp-in-2022-minister/

Ngandee, S., Taparugssanagorn, A., Anutariya, C., & Kuwornu, J. K. (2021). Assessment of rice yield prediction models based on big data analytics for better supply chain decision-making in Thailand. *International Journal of Value Chain Management*, *12*(3), 221-240.

Nkechi, I. (2022, September 16). Agriculture Contributed 23.3% to National GDP in Q2. *In Science Nigeria.* Retrieved December 20, 2024, from https://sciencenigeria.com/agriculture-contributed-23-3-to-national-gdp-in-q2-abubakar/

Okwuchi, I. (2020). Machine learning based models for fresh produce yield and price forecasting for strawberry fruit (Master's thesis, University of Waterloo).

Olayide, O. E., Tetteh, I. K., & Popoola, L. (2016). Differential impacts of rainfall and irrigation on agricultural production in Nigeria: Any lessons for climate-smart agriculture? Agricultural water management, 178, 30-36.

Padmasiri, G. A. S. M. (2022). Coconut Price Prediction in Sri Lanka Using Supervised Machine Learning Approach (LSTM) (Doctoral dissertation).

Pinheiro, C. and Senna, V. (2017). Multivariate analysis and neural networks application to price forecasting in the Brazilian agricultural market. Ciência Rural, 47(1). https://doi.org/10.1590/0103-8478cr20160077

Rakhra, M., Sanober, S., Quadri, N. N., Verma, N., Ray, S., & Asenso, E. (2022). Implementing machine learning for smart farming to forecast farmers' interest in hiring equipment. Journal of Food Quality, 2022.

RL, M., & Mishra, A. K. (2021). Forecasting spot prices of agricultural commodities in India: Application of deep-learning models. Intelligent Systems in Accounting, Finance and Management, 28(1), 72-83.

Shimizu, G. D., & Gonçalves, L. S. A. (2023). AgroReg: Main regression models in agricultural sciences implemented as an R Package. Scientia Agricola, 80, e20220041.

Singh, K., Singh, K., Kumar, S., Panwar, S., & Gurung, B. (2019). Forecasting crop yield through weather indices through lasso. The Indian Journal of Agricultural Sciences, 89(3). https://doi.org/10.56093/ijas.v89i3.87602

Singh, R., & Deorari, R. (2022, October). A Multiple Linear Regression Model for Crop Production using Machine Learning and Neural Network. In 2022 IEEE 2nd Mysore Sub sSection International Conference (MysuruCon) (pp. 1-6). IEEE.

Snee, R. D. (1977). Validation of regression models: methods and examples. Technometrics, 19(4), 415-428.

Statista. (n .d.). Agriculture in Nigeria- statistics and facts. *In Statista.* Retrieved December 20, 204, from https://www.statista.com/topics/6729/agriculture-in-nigeria/#topicOverview

Sun, F., Meng, X., Zhang, Y., Wang, Y., Jiang, H., & Liu, P. (2023). Agricultural Product Price Forecasting Methods: A Review. Agriculture, 13(9), 1671.

The World Bank (2020). Employment in Agriculture (% of Total Employment) (Modeled ILO Estimate) – Nigeria. *In The World Bank.* Retrieved January 20, 2024, from https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?locations=NG

Ukpe, W. (2023). GDP: Nigeria's Agriculture sector grows by 1.20% in Q2 2022. *In Nairametrics.* Retrieved December 20, 2024, from https://nairametrics.com/2022/08/26/gdp-nigerias-agriculture-sector-grows-by-1-20-in-q2-2022/

Verter, N., & Bečvářová, V. (2015). An analysis of yam production in Nigeria. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 63(2), 659-665.

Vikranth, K., S., N., & K., K. (2022). A survey of data analytical techniques on commodity price prediction for smart agricultural systems. International Journal of Case Studies in Business IT and Education, 259-267. https://doi.org/10.47992/ijcsbe.2581.6942.0194

Wang, L., Feng, J., Sui, X., Chu, X., & Mu, W. (2020). Agricultural product price forecasting methods: research advances and trends. British Food Journal, 122(7), 2121-2138. https://doi.org/10.1108/bfj-09-2019-0683

Wang, Y. (2023). Agricultural products price prediction based on improved RBF neural network model. Applied Artificial Intelligence, 37(1), 2204600.

Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M., & Wang, F. (2019). Forecasting horticultural products price using the Arima model and neural network based on a large-scale data set collected by the web crawler. IEEE Transactions on Computational Social Systems, 6(3), 547-553. https://doi.org/10.1109/tcss.2019.2914499

Wihartiko, F. D., Nurdiati, S., Buono, A., & Santosa, E. (2021, March). Agricultural price prediction models: a systematic literature review. In Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore (pp. 2927-2934).

Xie, H. and Wang, B. (2017). An empirical analysis of the impact of agricultural product price fluctuations on China's grain yield. Sustainability, 9(6), 906. https://doi.org/10.3390/su9060906

Zhang, Y. and Na, S. (2018). A novel agricultural commodity price forecasting model based on fuzzy information granulation and mea-svm model. Mathematical Problems in Engineering, 2018, 1-10. https://doi.org/10.1155/2018/2540681

Zhang, Z., Brizmohun, R., Li, G., & Wang, P. (2022). Does economic policy uncertainty undermine the stability of agricultural imports? evidence from China. Plos One, 17(3), e0265279. https://doi.org/10.1371/journal.pone.0265279