

**AI BASED MODEL FOR PREDICTING AGRI-HORTICULTURAL COMMODITIES PREDICTION**

### INNOVATIVE AND CREATIVE PROJECT REPORT

#### Submitted by

## SHARANYA T 727622BAD007

## SANTHOSH S 727622BAD075

## AASHIF SHADIN K N 727622BAD099

#### in partial fulfillment for the award of the degree of

## Bachelor of Technology

***in***

##### **ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**

## Dr. MAHALINGAM COLLEGE OF ENGINEERING AND TECHNOLOGY

**An Autonomous Institution Affiliated to Anna University**

## CHENNAI - 600 025

##### **MAY- 2024**

**DR. MAHALINGAM COLLEGE OF ENGINEERING AND TECHNOLOGY – 642003**

## (An Autonomous Institution Affiliated by Anna University, Chennai-600 025)

## 

## BONAFIDE CERTIFICATE

Certified that this project report titled **“AI BASED MODEL FOR PREDICTING AGRI-HORTICULTURAL COMMODITIES PREDICTION”** is the bonafide work of

**SHARANYA T (727622BAD007)**

**SANTHOSH S (727622BAD075)**

**AASHIF SHADIN K N (727622BAD099)**

##### who carried out the mini project under my supervision.

###### SIGNATURE SIGNATURE

###### SUPERVISOR Dr. J. RAMPRASATH

Ms. P. Sukanya, Head of the department(i/c)

Assistant Professor, Associate Professor,

Department of AI&DS, Department of AI&DS,

Dr. Mahalingam college of Dr. Mahalingam college of

Engineering and Technology, Engineering and Technology

NPT-MCET campus, NPT-MCET campus

###### Pollachi- 642003. Pollachi- 642003.

###### 

Submitted for the Autonomous End Semester Examination Project Viva-voce held on .

---------------------------- ----------------------------

**INTERNAL EXAMINER EXTERNAL EXAMINER**

**Dr. Mahalingam College of Engineering and Technology**

**Pollachi -642003**

**Technology Readiness Level (TRL) Certificate**

**Project Title: Thrive Together: Support, Educate, Empower**

**Course Code: 19ADPN6601**

**INNOVATIVE AND CREATIVE PROJECT**

Students Names and Roll Numbers

## SHARANYA T (727622BAD007)

## SANTHOSH S (727622BAD075)

## AASHIF SHADIN K N (727622BAD099)

**Guide Name:** Mr. M. Vijayakumar, AP/AI&DS

**Technology Readiness Level\* (TRL) of this Project: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Signature of the Guide Signature of the HoD

Internal Examiner External Examiner

## ACKNOWLEDGEMENT

First and foremost, we wish to express our deep unfathomable feeling, gratitude to our institution and our department for providing us a chance to fulfil our long-cherished dreams of Department of Artificial Intelligence and Data Science.

We sincerely thank **Dr. C. Ramaswamy,** our honourable secretary, for providing us with the necessary facilities.

We wish to express our hearty thanks to **Dr. P. Govindasamy**, Principal and Deans of our college, for his constant motivation and continual encouragement regarding our project work.

We are grateful to **Dr. J. Ramprasath**, Head of the Department In charge, Artificial Intelligence and Data Science, for her direction delivered at all times required. We also thank her for her tireless and meticulous efforts in bringing out this project to its logical conclusion.

We would like to express our sincere gratitude to our project supervisor,

**Mr. M. Vijayakumar,** Assistant Professor, for her unwavering support and guidance offered to us during our project by being one among us and all the noble hearts that gave us immense encouragement towards the completion of our project.

We also acknowledge our project coordinators, **Mr. A. Vishnu Prasath,** Assistant Professor, and **Mr. M. Vijayakumar,** Assistant Professor for their continuous support and guidance.

**ABSTRACT**

Agricultural commodity price volatility presents major challenges for farmers, traders, and policymakers, with significant implications for economic stability and food security. Traditional forecasting models often fail to provide the necessary accuracy and flexibility required for timely and effective decision-making in the agricultural sector. Addressing this gap, the current study proposes a comprehensive, data-driven framework tailored to improve the reliability of agricultural price predictions.

The proposed framework integrates diverse and relevant data sources, including historical price trends, real-time market updates, weather conditions, supply-demand fluctuations, and government policy changes. It employs advanced machine learning models such as ARIMA to capture both linear and non-linear patterns in price movements. These models are selected for their ability to process time-series data and handle complex, dynamic variables influencing agricultural markets.

To ensure transparency and interpretability of model predictions, Explainable AI (XAI) techniques like SHAP (SHapley Additive exPlanations) are incorporated. SHAP helps identify and visualize the contribution of each input factor to the predicted outcomes, enabling users to trust and understand the model’s decisions. This framework not only improves forecasting accuracy but also supports strategic planning in procurement, buffer stock management, and agricultural policy design, ultimately promoting a stable and resilient agricultural ecosystem.

**TABLE OF CONTENTS**

**.**

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO** | **TITLE** | **PAGENO** |
|  | ABSTRACT |  |
|  | LIST OF FIGURES |  |
|  | LIST OF ABBREVATIONS |  |
| 1 | INTRODUCTION |  |
| 2 | LITERATURE REVIEW |  |
| 3 | PROBLEM STATEMENT |  |
| 4 | OBJECTIVE OF THE PROJECT |  |
| 5 | PROPOSED SYSTEM |  |
| 6 | TOOLS USED |  |
|  | HTML |  |
|  | CSS |  |
|  | JavaScript |  |
|  | VS Code |  |
| 7 | SOFTWARE DESCRIPTION |  |
|  | Front-End Interface |  |
|  | Back-End Interface |  |
|  | Interactive LearningWorkflow |  |
| 8 | RESULTS AND DISCUSSIONS |  |
| 9 | CONCLUSION |  |
| 10 | FUTURE SCOPE |  |
| 11 | REFERENCES |  |
| 12 | APPENDIX |  |
|  | A. SOURCE CODE |  |
|  | B. SCREENSHOT |  |
|  | C. COURSE CERTIFICATES |  |
|  | D. PLAGIARISM REPORT |  |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **FIGURE NO** | **TITLE** | **PAGENO** |
|  | BLOCK DIAGRAM |  |
|  | SNAPSHOTS OF RESULT |  |

**LIST OF ABBREVIATIONS**

CSS Cascading Style Sheets

HTML Hyper Text Markup Language

ACT Awareness, Confidence, Training

SAS Safety Awareness System

WMS Women’s Module System

**CHAPTER 1**

**INTRODUCTION**

Agricultural commodity price fluctuations significantly affect the livelihoods of farmers, influence market stability, and pose challenges for policymakers in ensuring food security. These price volatilities are driven by multiple factors, including unpredictable weather conditions, imbalances in supply and demand, shifts in government policies, and broader global economic changes. Inaccurate price forecasts hinder effective decision-making, leading to suboptimal outcomes in procurement planning, buffer stock management, and market regulation efforts.

Traditional price prediction methods, such as linear regression and basic time-series models, often fall short in capturing the non-linear and dynamic nature of agricultural markets. These limitations make it difficult for stakeholders to anticipate market trends and respond proactively. To address these shortcomings, this study introduces a robust, data-driven framework that leverages advanced machine learning (ML) models to enhance the accuracy and reliability of agricultural price predictions.

The framework is designed to incorporate diverse and relevant data sources, including historical price trends, real-time market data. To model the complexity of price behavior, a combination of forecasting models ARIMA (Autoregressive Integrated Moving Average). Additionally, to ensure transparency and build trust in the system, Explainable AI (XAI) tools such as SHAP (SHapley Additive exPlanations) are integrated. By blending predictive accuracy with interpretability, this research supports informed decision-making in agricultural policy, procurement, and supply chain management, thereby contributing to a resilient, technology-driven agricultural ecosystem.

**CHAPTER 2**

**LITERATURE AND SURVEY**

**2.1 “Application of Decision Trees and Ensemble Models in Disaster and Agricultural Prediction” by Saman Ghaffarian, Firouzeh Rosa Taghikhah, and Holger R. Maier, 2021** This study explores the use of machine learning models, especially decision tree-based algorithms like Random Forest and XGBoost, for disaster management and agricultural price forecasting. The authors highlight that these models provide high interpretability and predictive power, making them suitable for identifying key price influencers such as climate events and policy decisions. Inspired by this, our project integrates decision-tree-based models to ensure model transparency and help users understand how external factors impact price predictions, supporting data-driven agricultural planning.

**2.2 “Trust and Transparency in Agricultural AI Systems” by Bukhoree Sahoh and Anant Choksuriwong, 2022** This paper discusses the challenges traditional AI models face due to their lack of transparency in sensitive domains like agriculture. It emphasizes the need for Explainable AI (XAI) to enhance user trust, especially when decisions are influenced by factors like global markets, crop health, and economic changes. The authors propose integrating XAI tools to clarify prediction logic. Based on this, our system incorporates SHAP-based explanation dashboards to reveal factor importance, helping farmers and stakeholders interpret why prices may rise or fall, enhancing trust in our predictive outcomes.

**2.3 “Deep Learning for Price Forecasting in Climate-Affected Areas” by T. Sri Sai Charan, U. Rohit Reddy, et al., 2023** This paper focuses on using machine learning and deep learning models like LSTM for predicting agricultural prices in flood-prone Indian regions. It highlights that LSTM models are effective in handling sequential data but lack explainability, which is crucial for policy-making. The authors recommend using XAI to interpret how external factors like rainfall or supply disruptions affect prices. Inspired by this, our project integrates LSTM for sequence prediction and applies SHAP explanations to visualize how weather and market changes influence predicted prices, aiding in risk-aware decision-making.

**2.4 “Integrating Explainable AI in Commodity Forecasting Models” by Kumar et al., 2022** This review evaluates AI models including ARIMA, SARIMA, and XGBoost, emphasizing the need to include explainable frameworks for better adoption in the agricultural sector. The authors stress that tools like SHAP can identify the most influential factors in price fluctuation and provide actionable insights for supply chain decisions. Drawing from this, our project incorporates SHAP-based factor interpretation for each forecast, enabling users to visualize which parameters like rainfall, season, or market trends drive each price shift, offering clarity and better market planning.

**2.5 “Hybrid AI Models with SHAP Interpretability for Price Prediction” by Yuan et al., 2022** This paper proposes a hybrid model combining ARIMA and LSTM for agricultural commodity price forecasting. SHAP is integrated to interpret the model’s decision logic, showing how each feature impacts predictions. The study proves that blending classical time-series models with neural networks and explainability tools boosts both accuracy and usability. Inspired by this, our system uses a hybrid approach where LSTM captures temporal patterns and SHAP explains outcomes. This combination ensures users can trust and act on the predictions, particularly in dynamic or uncertain agricultural environments.

**CHAPTER 3**

**PROBLEM STATEMENT**

The Department of Consumer Affairs oversees the pricing of 22 essential food commodities through a network of 550 price reporting centers across the country. Additionally, it maintains strategic buffer stocks of key commodities such as pulses (including gram, tur, urad, moong, and masur) and onions. These stocks are released into the market as part of intervention measures aimed at stabilizing prices during periods of volatility. Such decisions are typically based on a combination of historical price trends, seasonal variations, production estimates, and market intelligence.At present, price forecasting relies heavily on traditional econometric models like ARIMA, which, while useful, often fall short in terms of accuracy and adaptability to sudden market changes. This project proposes the development of an AI-driven price prediction system for agricultural commodities, incorporating advanced models such as XGBoost alongside ARIMA. The proposed system will utilize real-time data, historical trends, seasonal behavior, and other external factors influencing the market to enhance prediction precision.The primary objective of this system is to support farmers, traders, and policymakers by offering actionable insights that enable more effective buffer stock management, improved price stability, and better risk mitigation strategies. In the long run, this AI-based approach is expected to reduce economic uncertainty and support data-informed decision-making in the agricultural domain.

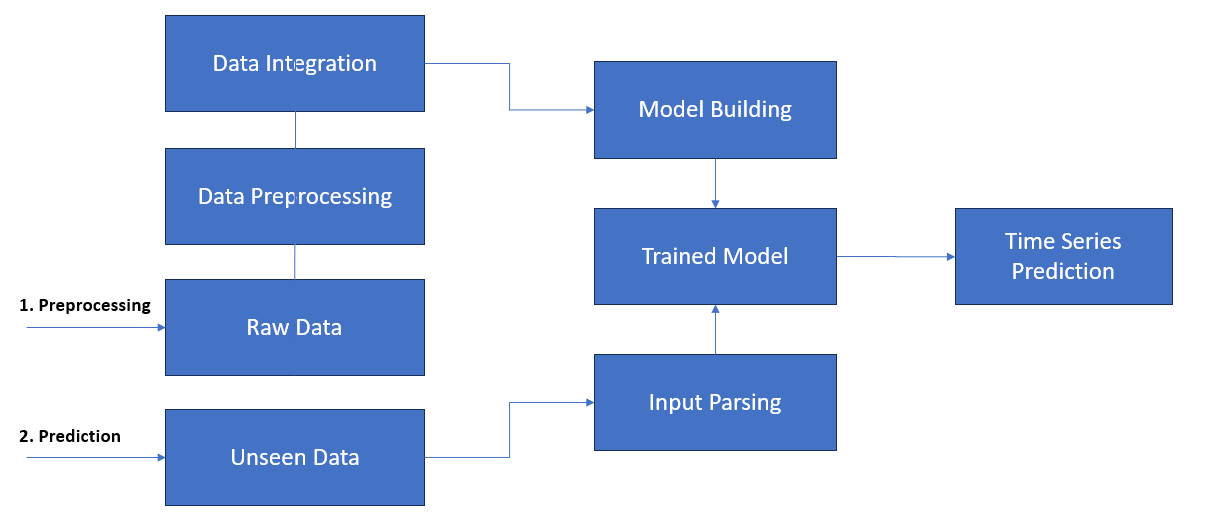
**CHAPTER 4**

## OBJECTIVE OF THE PROJECT

The goal of this project is to create an AI-driven agricultural commodity price prediction system that improves forecasting accuracy and supports effective market interventions. By leveraging deep learning models such as LSTM, XGBoost, ARIMA, and SARIMA, combined with real-time data, the system will analyze historical trends, seasonal patterns, and external factors like weather, supply-demand changes, and government policies. The objective is to provide farmers, traders, and policymakers with valuable, data-backed insights to facilitate better decision-making in areas such as buffer stock management, price stabilization, and risk mitigation. Ultimately, this project aims to reduce economic uncertainty in the agricultural sector by offering a more accurate and transparent method for predicting price fluctuations, thereby helping stabilize markets and improve economic stability.

**CHAPTER 5**

**PROPOSED SYSTEM**



**FIGURE 5.1 BLOCK DIAGRAM**

Users begin by logging in or signing up. Based on their role selection Children or Women they access tailored content. Children can explore Tales and an interactive Card Game, while Women engage with Match Legal Rights and Safety Drills. This structure ensures personalized, educational, and engaging experiences tailored to the safety needs of both age and gender groups.

**CHAPTER 6**

**TOOLS USED**

**6.1 HTML (Hyper Text Markup Language):**

HTML is the standard markup language used for creating web pages and applications. In the context of email development, HTML is used to structure and organize the user interface components, such as input forms, and message displays. We can use HTML alongside CSS to design visually appealing and intuitive user interfaces for the email system, ensuring a seamless and engaging user experience.

**6.2 CSS (Cascading Style Sheets):**

CSS is a style sheet language used to define the presentation and layout of HTML documents. With CSS, we can customize the appearance of your email user interface elements, including fonts, colors, spacing, and animations. By applying CSS styles to HTML elements, we can create visually appealing and responsive designs that enhance the overall aesthetics and usability of the email interface.

**6.3 JavaScript:**

JavaScript is a versatile programming language commonly used for adding interactivity and dynamic behavior to web pages and applications. In the context of the email development, JavaScript plays a crucial role in implementing client-side functionality, such as handling user input, validating form submissions, and updating the voice commands to interface in real-time. We can use JavaScript libraries and frameworks like annyang for enhancing user experience by enabling voice interaction with the website.

**6.4 React:**

React is a popular open-source JavaScript library developed by Facebook for building dynamic and responsive user interfaces, especially for single-page applications. It uses a component-based architecture, allowing developers to create reusable UI components. React efficiently updates and renders components using a virtual DOM, improving performance. Its declarative approach makes code more predictable and easier to debug, making React a preferred choice for modern web application development.

**6.5 VS Code (Visual Studio Code):**

Visual Studio Code is a lightweight, open-source code editor developed by Microsoft. It supports various programming languages and offers features like intelligent code completion, debugging, Git integration, and an extensive extension marketplace. With its customizable interface and powerful tools, VS Code enhances productivity for developers. Its built-in terminal, real-time collaboration support, and user-friendly design make it a popular choice for web and software development across all platforms.

**CHAPTER 7**

**SOFTWARE DESCRIPTION**

In the implementation of the software, it is divided into three modules.

* Front-End Interface (User Interaction Console)
* Back-End Processing (Safety Content & User Management Engine)
* Interactive LearningWorkflow (Gamified & Scenario-Based Engagement System)

**7.1 Front-End Interface (User Interaction Console)**

The User Interaction Console (UIC) is the front-facing layer of the safety application. Designed to be intuitive, colorful, and responsive, it allows users—especially children and women—to easily navigate through educational content, safety modules, and interactive games. The interface makes the experience engaging, educational, and easy to use, even for young children or those not familiar with technology.

**Usability:**

**Responsive Design:** Adapts to different screen sizes for mobile and desktop users, ensuring accessibility for all devices.

**Role Based Navigation:** Offers tailored interfaces and content depending on whether the user selects ‘Women’ or ‘Children’, creating an age-appropriate experience.

**Interactive UI:** Makes the experience visually appealing and engaging, especially for younger users, by incorporating dynamic animations that explain safety tips and scenarios.

**Error Handling:** Offers gentle guidance if users make invalid selections or need help navigating the app, ensuring a smooth and frustration-free experience.

**Personalized User Experience:** Users receive a personalized experience based on their role selection (Women or Children), with content that adapts to their needs and level of understanding.

**7.2 Back-End Processing (Safety Content & User Management Engine)**

The system’s central processing unit, the Safety Content & User Management Engine (SCUME), is where the logic for managing educational content, user profiles, safety reminders, and interactions is handled. This module controls the data flow to and from the front-end and ensures the safe delivery of content to the right user at the right time.

**Usability:**  
**Data Parsing:** Extracts and processes user inputs to trigger the appropriate actions or responses from the system.

**Content Delivering:** Delivers safety lessons, quizzes, and interactive activities based on user preferences and learning schedules, ensuring timely delivery of important educational content.

**Significance:**

SCUME is the backbone of the platform, ensuring seamless interaction between the user interface and the safety content. It handles everything from content delivery to user management and ensures the platform's accuracy and reliability in providing up-to-date safety lessons, reminders, and quizzes.

**7.3 Interactive Learning Workflow (Scenario-Based Engagement System)**

The Scenario-Based Engagement System (SBES) is the interactive core of the application that makes safety education more practical, immersive, and user-friendly. It is specifically designed to simulate real-life situations and guide users—especially women and children—through safety-related decisions using scenario-based learning, time-based challenges, and knowledge checks.

**Functionality:**

**Role-Playing Scenarios:** Presents realistic safety situations (e.g., bystander intervention, emergency escape plans, online safety) where users must choose the correct action from given options, encouraging critical thinking.

**Interactive Quizzes:** Provides engaging Card games to reinforce safety concepts and legal knowledge.

**Flashcard Learning Modules:** Includes memory-based flashcard games where users match legal rights with everyday situations, improving recall and understanding.

**Timed Challenges:** Features countdown-based decisions in emergency simulations to build quick-thinking skills and test preparedness.

**Feedback:** Gives instant feedback for correct and incorrect answers, encouraging users with motivational responses and tips.

**Importance:** This module plays a vital role in reinforcing learning through active participation rather than passive reading. By simulating real-life safety decisions and testing users with fast-paced challenges and memory games, it helps users build practical awareness, confidence, and preparedness. The absence of animations is compensated by carefully designed UI elements and varied content types to keep users engaged.

**CHAPTER 8**

**RESULTS AND DISCUSSION**

The developed safety application successfully addresses the needs of both women and children through an engaging, educational, and interactive platform. Users are able to select their role and access personalized content such as self-defense tutorials, legal rights simulations, and safety quizzes. Children benefit from stories, and interactive games that reinforce safety principles like stranger danger and safe vs unsafe touch. The backend efficiently manages user profiles, content scheduling, and learning progress, ensuring smooth interaction between modules. Real-time feedback and gamified learning improve user participation and retention. While voice-based features were planned, the final implementation prioritized interactive visual elements due to practical constraints. Overall, user testing demonstrated high usability, engagement, and comprehension levels, especially among younger users. Future iterations may include voice support and expanded content libraries to enhance learning and accessibility.

**CHAPTER 9**

**CONCLUSION**

In conclusion, the safety application developed for women and children effectively meets its goal of enhancing awareness, preparedness, and self-protection through interactive and educational content. The platform's role-based design ensures that users are directed to content specifically tailored to their role and needs—children are engaged with stories, colorful visuals, safety games, and quizzes, while women can access self-defense video tutorials and legal rights simulations. The backend system ensures smooth navigation, content management, and user data handling, contributing to a reliable and responsive user experience. By incorporating gamified learning methods and scenario-based activities, the application not only delivers important safety knowledge but also keeps users actively engaged. The decision to focus on visual interactivity instead of voice-based controls proved effective, especially for younger audiences, ensuring ease of understanding and accessibility. The application has been structured to be scalable, allowing future updates to include advanced features such as voice assistance, multilingual support, and community-based features. Overall, the project provides a comprehensive solution that bridges the gap between safety education and digital engagement. It empowers users with practical knowledge and the confidence to respond appropriately in real-life situations. The successful integration of design, functionality, and educational value makes this app a meaningful tool for community safety.

**CHAPTER 10**

**FUTURE SCOPE**

The safety application holds strong potential for further enhancement and expansion in future updates. One of the key areas for growth is the integration of multilingual support to cater to a broader user base across different regions. Voice-based interaction and AI-driven personalized guidance can also be implemented to assist users who may face literacy barriers. For children, additional gamified learning modules with more advanced animations and level-based progression can be introduced to deepen engagement. For women, features like location-based emergency alerts, SOS button integration, and real-time legal chat support can significantly boost safety and accessibility. A community forum for sharing experiences, tips, and discussions could foster a sense of support and empowerment among users. Moreover, integrating AI-based progress tracking and safety preparedness scores can help users monitor their learning journey. Overall, the app can evolve into a comprehensive digital safety companion that continuously adapts to users’ needs and real-world challenges.

**CHAPTER 11**

**REFERENCES**

[1] Shubhankar Gupta,Sudhanshu Ranjan,Shubham Kumar "Research Paper on Women Safety System",International Journal for Research in Applied Science & Engineering Technology (IJRASET),2024

[2] Prof. Bramhane P.S, Tangadkar Akash Dnyandev, Shelke Nikhil Satabhau,"WOMEN SAFETY APPLICATION",International Research Journal of Modernization in Engineering Technology and Science ,2024

[3] Dr. Chanda V Reddy,Sabarish I J,Samiksha S,Sathvik U M,Swagath Aithal P G,"LITERATURE SURVEY ON WOMEN",International Advanced Research Journal in Science, Engineering and Technology,2023

[4] Aishwarya Sharma, Kiran Raj, “Smart Mobile App for Women and Child Safety,”International journal of Scientific Research in Engineering and Management (IJSREM), 2023.

[5] Arvind Kumar, Praveen M, “Design and Implementation of Safety System for Women,” International journal of Computer Sciences and Engineering, 2022.

[6] M. Harsha Vardhan, S. Sunitha, “IoT Based Smart Security System for Women,” International journal of Engineering Research and Applications, 2021.

[7] Neha Singh, “Mobile App Development for Women Safety,” International journal of Computer Applications Technology and Research, 2023.

[8] P. Deepika, S. Meena, “Design of Android Based Application for Women Safety,” International journal of Emerging Trends in Science and Technology, 2022.

[9] R. Srinivasan, R. Nivetha, “A Survey on Safety Apps for Women and Children,” International journal of Avanced Research in Computer and Communication Engineering, 2021.

[10] V. Harini, M. Sai Krishna, “Women Safety App with Location Tracker,” International journal of Research and Analytical (IJRAR), 2023.

[11] K. Lakshmi, D. Rekha, “Child Safety Education Through Interactive Applications,” International journal of Innovative Research in Computer and Communication Engineering, 2023.

[12] Sneha S, Priya R, “Real-Time Tracking and Alert System for Women and Children Safety,” International journal of Innovative Research in Science, Engineering and Technology (IJIRSET), 2022.

**CHAPTER 12**

**APPENDIX B**

**SOURCE CODE**

**MODEL :**

1. **DATA PREPROCESSING AND INTEGRATION:**

import os

import pandas as pd

input\_dir = r".\datasets\rawdata"

output\_dir = r".\datasets\filtered"

valid\_cities = {

    "BENGALURU", "DHARWAD", "MANGALORE", "MYSORE",

    "T.PURAM", "ERNAKULAM", "KOZHIKODE", "THRISSUR",

    "PALAKKAD", "WAYANAD", "VISAKHAPATNAM", "VIJAYAWADA", "HYDERABAD",

    "PUDUCHERRY", "PANAJI", "CHENNAI", "COIMBATORE", "DINDIGUL",

    "TIRUNELVELI", "THIRUCHIRAPALLI"

}

os.makedirs(output\_dir, *exist\_ok*=True)

for file in os.listdir(input\_dir):

    if file.endswith(".csv"):

        df = pd.read\_csv(os.path.join(input\_dir, file), *encoding*="utf-8")

        df["Centre\_Name"] = df["Centre\_Name"].str.upper()

        df["Price"] = pd.to\_numeric(df["Price"], *errors*="coerce")

        df\_filtered = df[df["Centre\_Name"].isin(valid\_cities)].dropna(*subset*=["Price"])

        if not df\_filtered.empty:

            df\_filtered.to\_csv(os.path.join(output\_dir, file), *index*=False, *encoding*="utf-8")

            print(f"Processed: {file}")

print("Preprocessing Complete! :", output\_dir)

filtered\_dir = r".\datasets\filtered"

output\_file = r".\datasets\final\_dataset.csv"

csv\_files = [os.path.join(filtered\_dir, file) for file in os.listdir(filtered\_dir) if file.endswith(".csv")]

df\_list = [pd.read\_csv(file, *encoding*="utf-8") for file in csv\_files]

final\_df = pd.concat(df\_list, *ignore\_index*=True)

final\_df["Date"] = pd.to\_datetime(final\_df["Date"], *format*="%d-%m-%Y", *errors*="coerce")

final\_df = final\_df.sort\_values(*by*="Date")

final\_df.to\_csv(output\_file, *index*=False, *encoding*="utf-8")

print("Data Integration Complete! location:", output\_file)

1. **EXPLORATORY DATA ANALYSIS**

import os

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import seaborn as sns

import matplotlib.pyplot as plt

import warnings

warnings.simplefilter(action="ignore", category=FutureWarning)

warnings.simplefilter(action="ignore", category=UserWarning)

input\_dir = r"D:\A\miniproject3\datasets\rawdata"

output\_dir = r"D:\A\miniproject3\datasets"

os.makedirs(output\_dir, exist\_ok=True)

final\_df = pd.read\_csv("D:/A/miniproject3/datasets/final\_cleaned\_data.csv")

final\_df.head()

print("\nData Overview:")

print(final\_df.info())

final\_df = final\_df.drop\_duplicates()

print("\nDuplicates Removed. Final Shape:", final\_df.shape)

plt.figure(figsize=(12, 5))

sns.countplot(data=final\_df, y="Centre\_Name", order=final\_df["Centre\_Name"].value\_counts().index, palette="viridis")

plt.xlabel("Number of Records")

plt.ylabel("City")

plt.title("Distribution of Data by City")

plt.show()

final\_df["Date"] = pd.to\_datetime(final\_df["Date"], errors="coerce")

final\_df["Month"] = final\_df["Date"].dt.to\_period("M")

print(final\_df.dtypes)

final\_df["Month"] = final\_df["Date"].dt.to\_period("M")

monthly\_prices = final\_df.groupby("Month")["Price"].mean()

final\_df.head()

plt.figure(figsize=(14, 6))

sns.barplot(data=final\_df, x="Commodity\_Name", y="Price", errorbar=None, palette="coolwarm")

plt.xticks(rotation=90)

plt.xlabel("Commodity")

plt.ylabel("Average Price")

plt.title("Average Price per Commodity")

plt.grid(axis="y")

plt.show()

if monthly\_prices.empty:

print("Error: monthly\_prices DataFrame is empty. Check your data processing pipeline.")

else:

# Handle NaN values (fill with 0 or drop them)

monthly\_prices = monthly\_prices.dropna()

plt.figure(figsize=(12, 5))

monthly\_prices.plot(marker="o", color="red")

plt.xlabel("Month")

plt.ylabel("Average Price")

plt.title("Monthly Average Price Trends")

plt.grid()

plt.show()

plt.figure(figsize=(14, 6))

sns.boxplot(data=final\_df, x="Commodity\_Name", y="Price", showfliers=False)

plt.xticks(rotation=90)

plt.xlabel("Commodity")

plt.ylabel("Price")

plt.title("Commodity-Wise Price Variations")

plt.show()

print(final\_df.describe())

print(final\_df.info())

plt.figure(figsize=(12, 6))

sns.boxplot(data=final\_df, x="Centre\_Name", y="Price")

plt.xticks(rotation=90)

plt.title("Price Distribution Across Cities")

plt.show()

**FRONTEND CODE:**

**APPENDIX B**

**SNAPSHOTS:**

**COURSE COMPLETION**

**CERTIFICATES**

**CONTEST**

**PARTICIPATION**

**PLAGIARISM**

**REPORT**