VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI – 590018, Karnataka

INTERNSHIP REPORT

ON

"A predictive model for forecasting demand and supply information of TOP crops"

Submitted in partial fulfilment for the award of degree(18ETI85)

BACHELOR OF ENGINEERING IN ELECTRONICS AND TELECOMMUNICATION ENGINEERING

Submitted by:

ULLAS B

(1BY20ET061)

Conducted at

Compsoft Technologies

BMS INSTITUTE OF TECHNOLOGY AND MANAGEMENT

Avalahalli, Yelahanka, Bengaluru-560064

2023-24

BMS INSTITUTE OF TECHNOLOGY AND MANAGEMENT

Avalahalli, Yelahanka, Bengaluru-560064

2022-23

CERTIFICATE

This is to certify that the Internship titled "A predictive model for forecasting demand and supply information of TOP crops" carried out by Mr. ULLAS B (1BY20ET061) a bonafide student of BMS Institute of Technology & Management, in partial fulfillment for the award of Bachelor of Engineering, in ELECTRONICS AND TELECOMMUNICATION ENGINEERING under Visvesvaraya Technological University, Belagavi, during the year 2022-2023. It is certified that all corrections/suggestions indicated have been incorporated in the report.

The project report has been approved as it satisfies the academic requirements in respectof Internship prescribed for the course Internship / Professional Practice (18ETI85)

Signature of Guide **Signature of HOD** **Signature of Principal**

External Viva:

Name of the Examiner Signature with Date

1)_____

_____

2)_____

_____

D E C L A R A T I O N

I, **Ullas B**, final year student of Electronics and Telecommunication, BMSIT - 560 064, declare that the Internship has been successfully completed, in **Compsoft Technologies**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Electronics and Telecommunication, during the academic year 2022-2023.

Date :21-09-2023 :

Place : Bengaluru

USN : 1BY20ET061

NAME : ULLAS B

OFFER LETTER

Internship Offer Letter

Date: 11th August, 2023

Name: **Ullas B**
USN: **1BY20ET061**
Placement ID: **TIE0408ML301**

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning with Python (Research Based)** Internship position with **Compsoft Technologies**, effective Start Date **11th August, 2023**, All of us are excited about this opportunity provided to you!

This internship is viewed as being an educational opportunity for you, rather than a part-time job. As such, your internship will include training/orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts of **Machine Learning with Python (Research Based)** through hands-on application of the knowledge you learn while you train with the senior developers. You will be bound to follow the rules and regulations of the company during your internship duration.

Again, congratulations and we look forward to working with you!.

Sincerely,

Nithin K. S
Project Manager
Compsoft Technologies
No. 363, 19th main road,
1st Block Rajajinagar
Bangalore - 560010

A C K N O W L E D G E M E N T

This Internship is a result of accumulated guidance, direction and support of several important persons. We take this opportunity to express our gratitude to all who have helped us to complete the Internship.

We express our sincere thanks to our **Dr. Mohan Babu G N,** principal, BMSIT&M for providing usadequate facilities to undertake this Internship.

We would like to thank Dr. Mallikarjuna Gowda C P ,HoD, Deparment of Electronics and Telecommunication Engineering, BMSIT&M for providing us an opportunity to carry out Internship and for his valuable guidance and support.

We express our deep and profound gratitude to our guide, Dr. Mallikarjuna Gowda C P ,HoD, Deparment of Electronics and Telecommunication Engineering, BMSIT&M for his keen interest and encouragement at every step in completing the Internship.

We would like to thank all the faculty members of our department for the support extended during the course of Internship.

We would like to thank the non-teaching members of our dept, for helping us during the Internship.

Last but not the least, we would like to thank our parents and friends without whose constant help, the completion of Internship would have not been possible.

ULLAS B
(1BY20ET061)

ABSTRACT

Agriculture is a cornerstone of global food security and economic stability. Accurate forecasting of demand and supply information for top crops is essential for ensuring efficient resource allocation, mitigating food shortages, and supporting informed decision-making in the agricultural sector. This research presents a comprehensive predictive model designed to forecast the demand and supply information of top crops, addressing the critical need for reliable agricultural planning and management.

Our approach combines advanced data analytics, machine learning techniques, and domain-specific knowledge to create a robust predictive model. We collect and integrate diverse datasets, including historical crop production records, climate data, market trends, and demographic information. By leveraging this rich dataset, our model employs sophisticated algorithms to identify patterns, correlations, and trends that influence crop demand and supply dynamics.

Key components of our predictive model include:

Crop Yield Prediction: Utilizing historical data, climate information, and soil quality indicators, we forecast crop yields for the upcoming growing seasons. This enables farmers and policymakers to anticipate potential surpluses or shortages.

Demand Forecasting: We employ machine learning algorithms to predict future crop demand based on factors such as population growth, dietary trends, and economic indicators. This aspect helps stakeholders adjust their production strategies accordingly.

Supply Chain Analysis: Our model assesses supply chain factors, including transportation infrastructure, storage capacity, and distribution networks, to identify potential bottlenecks or inefficiencies in the agricultural supply chain.

Market Price Prediction: We integrate market data to forecast crop prices, allowing farmers to make informed planting decisions and enabling policymakers to develop effective price stabilization mechanisms.

The validation of our model is based on historical data and real-world case studies, demonstrating its accuracy and reliability in forecasting crop demand and supply dynamics. By providing timely and actionable insights, this predictive model contributes to sustainable agriculture, food security, and economic stability at local, regional, and global levels.

In conclusion, our research offers a holistic approach to addressing the challenges associated with forecasting demand and supply information for top crops. By leveraging data-driven insights and cutting-edge technology, this predictive model empowers stakeholders to make informed decisions, ultimately fostering a more resilient and efficient agricultural ecosystem

Table of Contents

CHAPTER 1

COMPANY PROFILE

1. COMPANY PROFILE

A Brief History of Company

Company, was incorporated with a goal ”To provide high quality and optimal Technological Solutions to business requirements of our clients”. Every business is a different and has a unique business model and so are the technological requirements. They understand this and hence the solutions provided to these requirements are different as well. They focus on clients requirements and provide them with tailor made technological solutions.They also understand that Reach of their Product to its targeted market or the automation of the existing process into e-client and simple process are the key features that our clients desire from Technological Solution they are looking for and these are the features that we focus on while designing the solutions for their clients.

Company is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Sarvamoola Software Services. specialize in ERP, Connectivity, SEO Services, Conference Management, effective webpromotion and tailor-made software products, designing solutions best suiting clients requirements.

we strive to be the front runner in creativity and innovation in software development through their well-researched expertise and establish it as an out of the box software development company in Bangalore, India. As a software development company, they translate this software development expertise into value for their customers through their professional solutions.

They understand that the best desired output can be achieved only by understanding the clients demand better. At our Company we work with them clients and help them todefiine their exact solution requirement. Sometimes even they wonder that they have completely redefined their solution or new application requirement during the brainstorming session, and here they position themselves as an IT solutions consulting group comprising of high caliber consultants.

They believe that Technology when used properly can help any business to scale and achieve new heights of success. It helps Improve its efficiency, profitability, reliability; to put itin one sentence ” Technology helps you to Delight your Customers” and that is what we wantto achieve.

CHAPTER 2

ABOUT THE COMPANY

2. ABOUT THE COMPANY

We are a Technology Organization providing solutions for all web design and development, Researching and Publishing Papers to ensure the quality of most used ML Models, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Compsoft Technologies specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements. The organization where they have a right mix of professionals as a stakeholders to help us serve our clients with best of our capability and with at par industry standards.They have young, enthusiastic, passionate and creative Professionals to develop technological innovations in the field of Mobile technologies, Web applications as well as Business and Enterprise solution. Motto of our organization is to "Collaborate with our clients to provide them with best Technological solution hence creating Good Present and Better Future for our client which will bring a cascading a positive effect in their business shape as well". Providing a Complete suite of technical solutions is not just our tag line, it is Our Vision for Our Clients and for Us, We strive hard to achieve it.

Services provided by Compsoft Technologies.

- Core Java and Advanced Java
- Research and Development/Improvise of ML Models
- Web services and development
- Dot Net Framework
- Python
- Selenium Testing
- Conference / Event Management Service
- Academic Project Guidance
- On The Job Training
- Software Training

CHAPTER 3

INTRODUCTION

3. INTRODUCTION

Introduction to ML

The agriculture industry is facing a pressing challenge in the 21st century: how to meet the increasing demand for food while simultaneously dealing with factors like climate change, limited resources, and changing consumer preferences. Accurate forecasting of demand and supply information for major crops is critical in addressing these challenges effectively. This is where Machine Learning (ML) plays a pivotal role.

Machine Learning, a subset of artificial intelligence, empowers computers to learn from historical data and make predictions or decisions without being explicitly programmed. In the context of agriculture and crop forecasting, ML techniques have emerged as powerful tools for optimizing production, resource allocation, and risk management.

In summary, Machine Learning has become an indispensable tool in the agriculture sector, providing the means to analyze vast and complex datasets, make accurate predictions, and support data-driven decision-making. This technology has the potential to revolutionize the way we approach crop forecasting, helping us address the challenges of food security and sustainability in a rapidly changing world. This research aims to explore and develop a predictive model harnessing the power of Machine Learning to forecast demand and supply information for major crops, contributing to more efficient and resilient agricultural systems.

Problem Statement

The current methods for forecasting demand and supply information of major crops in the agriculture sector are often limited in accuracy and efficiency. Traditional approaches lack the capability to process and analyze the vast and diverse datasets available, resulting in suboptimal decision-making, resource allocation, and risk management. Moreover, the dynamic nature of agriculture, influenced by factors like climate variability and market trends, demands a more adaptive and data-driven approach.

In this context, the problem statement is to develop and implement a Machine Learning-based predictive model that can:

- Accurately forecast crop yields based on historical and real-time data, incorporating variables such as weather conditions, soil quality, and crop-specific attributes.
- Predict future crop demand by analyzing a wide range of data sources, including demographic data, economic indicators, dietary trends, and market dynamics.
- Evaluate and mitigate risks associated with crop production, such as the impact of extreme weather events, pest outbreaks, and disease.

- Optimize the agricultural supply chain by analyzing transportation routes, storage capacities, and distribution networks to reduce waste and enhance efficiency.

- Provide real-time decision support to stakeholders, including farmers, policymakers, and food distributors, enabling them to make informed choices regarding crop management, resource allocation, and market strategies.

The goal is to leverage the power of Machine Learning to address these challenges comprehensively, ultimately contributing to a more resilient, sustainable, and efficient agricultural ecosystem. This research aims to design, develop, and validate a predictive model that can provide accurate, timely, and actionable insights into the demand and supply information of major crops, thereby enhancing food security and economic stability on a global scale.

CHAPTER 4

SYSTEM ANALYSIS

4. SYSTEM ANALYSIS

1. Existing System

The existing system for forecasting demand and supply information of major crops typically relies on a combination of traditional statistical methods, expert knowledge, and historical data analysis. While these methods have been used for many years and have provided valuable insights, they often have limitations that can be addressed by incorporating Machine Learning (ML) techniques. Here's an overview of the existing system:

Historical Data Analysis: The foundation of the existing system is historical data on crop production, yields, and prices. This data is collected over time and used to identify trends and patterns, which are then extrapolated to make predictions about future crop production and demand.

Expert Knowledge: Agriculture experts and researchers play a significant role in the existing system. They provide valuable insights and domain-specific knowledge to complement data analysis. Expert judgment is often used to adjust forecasts and make qualitative assessments.

Climate and Weather Data: Weather and climate data are critical inputs in the existing system. Historical weather data is analyzed to understand its impact on crop yields and to make predictions based on weather forecasts for upcoming seasons.

Market Analysis: Market analysts monitor market trends, consumer preferences, and economic indicators to anticipate changes in crop demand and prices. This analysis is often combined with historical market data to make forecasts.

Government Reports: Government agencies often release reports on crop production, supply, and demand. These reports provide valuable information for the existing system and are used by policymakers and stakeholders for decision-making.

Crop Modeling: Some existing systems use crop modeling techniques that simulate crop growth based on factors like soil quality, weather, and irrigation. These models can provide insights into potential crop yields.

While the existing system has been effective to some extent, it has several limitations:

Limited Scalability: Traditional methods struggle to efficiently process and analyze the increasingly large and complex datasets available today.

Inefficiency: Manual data collection and analysis processes can be time-consuming and prone to errors.

Limited Adaptability: Traditional methods may struggle to adapt to rapidly changing conditions, such as extreme weather events or emerging market trends.

Subjectivity: Expert judgment can introduce subjectivity and bias into forecasts.

Inadequate Risk Assessment: The existing system may not provide robust risk assessment and mitigation strategies for factors like climate change, pests, and diseases.

Lack of Real-time Insights: Timely decision-making can be hampered by the delayed availability of data and analysis results.

To address these limitations, there is a growing shift towards incorporating Machine Learning and advanced analytics into the existing system. ML models can handle large datasets, adapt to changing conditions, and provide more accurate and timely predictions for crop demand and supply, ultimately improving the efficiency and effectiveness of **agricultural decision-making processes.**

2. Proposed System

Proposed System for Machine Learning-Based Forecasting of Demand and Supply Information for Major Crops:

In response to the limitations of the existing system, we propose a modern and data-driven approach that leverages Machine Learning (ML) techniques to enhance the accuracy, scalability, and adaptability of forecasting demand and supply information for major crops. Our proposed system encompasses the following key components:

Data Integration and Preprocessing:

Comprehensive Data Sources: Gather diverse datasets, including historical crop production records, climate data, market trends, soil quality information, demographic data, and satellite imagery.
Data Cleaning and Preprocessing: Apply data cleaning, normalization, and feature engineering techniques to prepare the data for analysis.

Machine Learning Models:

Crop Yield Prediction: Develop ML models that can predict crop yields based on historical data, weather patterns, soil characteristics, and other relevant variables. Techniques such as regression, random forests, or neural networks can be employed.
Demand Forecasting: Utilize ML algorithms to forecast future crop demand by analyzing economic indicators, population growth, dietary trends, and market dynamics. Time series forecasting methods and deep learning models can be applied.
Risk Assessment: Develop ML models that assess and predict risks to crop production, including climate-related risks (e.g., droughts, floods), pest and disease outbreaks, and geopolitical factors. These models can provide early warnings and risk mitigation strategies.
Supply Chain Optimization: Employ ML algorithms to optimize the agricultural supply chain by analyzing transportation logistics, storage capacities, and distribution networks. Optimization techniques such as linear programming or reinforcement learning can be used.

Real-time Data Integration:

Integrate real-time data sources, including weather forecasts, market prices, and satellite imagery, to provide up-to-the-minute insights and improve the model's adaptability.

Decision Support System:

Develop a user-friendly decision support system (DSS) that integrates the ML models. The DSS should provide actionable recommendations to various stakeholders, including farmers, agricultural authorities, and food distributors.
Continuous Learning:

Implement a continuous learning framework that allows the ML models to adapt and improve over time as new data becomes available. Regular model retraining ensures accuracy and relevance.
Visualization and Reporting:

Create visualization tools and reports that communicate the model's findings effectively to end-users. Visualizations can help stakeholders understand trends, make informed decisions, and respond to emerging challenges.
Validation and Testing:

Rigorously validate the ML models using historical data and real-world case studies to ensure their accuracy and reliability.
Scalability and Accessibility:

Design the system to be scalable and accessible, ensuring that it can accommodate increasing data volumes and be used by stakeholders with varying levels of technical expertise.
The proposed system leverages the power of Machine Learning and data-driven insights to address the limitations of the existing system. It aims to provide more accurate, timely, and actionable information for forecasting demand and supply information for major crops, ultimately contributing to improved food security, sustainable agriculture, and economic stability.

3. Objective of the System

Objectives of the Machine Learning-Based System for Forecasting Demand and Supply Information for Major Crops:

Accurate Yield Prediction:

Develop machine learning models to accurately predict crop yields, taking into account historical data, weather patterns, soil quality, and crop-specific attributes.

Precise Demand Forecasting:

Utilize machine learning algorithms to provide precise forecasts of crop demand, incorporating economic indicators, population growth, dietary trends, and market dynamics.

CHAPTER 5

REQUIREMENT ANALYSIS

5. REQUIREMENT ANALYSIS

Hardware Requirement Specification

Hardware Requirement Specifications for a Predictive Model for Forecasting Demand and Supply Information of Top Crops:

Computing Infrastructure:

High-performance servers or cloud computing resources with multi-core processors to handle intensive data processing and machine learning tasks efficiently.
Consideration for scalability to accommodate growing data volumes and computational demands.

Memory (RAM):

Sufficient RAM to accommodate large datasets and facilitate faster data manipulation and model training. A minimum of 16 GB RAM is recommended, but more may be necessary for large-scale analysis.

Graphics Processing Unit (GPU):

GPUs, such as NVIDIA GPUs, for accelerating deep learning model training, especially if the predictive model involves complex neural networks.

Storage:

High-capacity storage for storing historical data, datasets, model checkpoints, and other relevant files.
Fast storage (e.g., SSDs) for quicker data access and model training.

Networking:

High-speed and reliable network infrastructure to facilitate data transfer, especially when working with large datasets and real-time data integration.

Security Measures:

Firewall systems and intrusion detection/prevention systems to protect against unauthorized access and cyber threats.
Encryption mechanisms for data security during storage and transmission.

Monitoring and Management Tools:

Hardware monitoring tools to track the health and performance of servers, storage, and network components.
Remote management capabilities for efficient system administration and troubleshooting.

User Interface Devices:

Desktop computers or workstations with adequate processing power and memory for data visualization, analysis, and model interpretation.
Input devices such as keyboards and mice, along with high-resolution monitors.

Budget Considerations:

Ensure that the hardware procurement aligns with the project budget constraints while providing the necessary computational resources to support the predictive modeling tasks.

The specific hardware requirements may vary depending on the scale of the project, the complexity of machine learning models used, and the size of the datasets involved. Regular maintenance, hardware upgrades, and performance monitoring should be part of the system's ongoing management to ensure its effectiveness in forecasting demand and supply information for top crops.

Software Requirement Specification

Software Requirement Specification for the Machine Learning-Based System for Forecasting Demand and Supply Information for Major Crops:

Operating System:

Linux-based server operating system (e.g., Ubuntu Server, CentOS) for hosting the backend infrastructure.
Windows or macOS for user interface devices (e.g., desktop computers).

Database Management:

Relational Database Management System (RDBMS) such as PostgreSQL or MySQL for storing structured data, historical records, and metadata.
NoSQL database (e.g., MongoDB) for handling unstructured or semi-structured data if required.

Programming Languages and Libraries:

Python as the primary programming language for developing machine learning models and data analysis scripts.
Python libraries such as NumPy, pandas, scikit-learn, TensorFlow, PyTorch, and Keras for machine learning model development.
Additional libraries for data manipulation, visualization (e.g., Matplotlib, Seaborn), and web development (e.g., Django, Flask) as needed.

Machine Learning Frameworks:

TensorFlow or PyTorch for building and training deep learning models, if applicable.
Scikit-learn for traditional machine learning algorithms.

CHAPTER 6

DESIGN ANALYSIS

6. DESIGN & ANALYSIS

Design and Analysis for the Machine Learning-Based System for Forecasting Demand and Supply Information for Major Crops involves planning the system's architecture, workflow, and evaluating the feasibility and effectiveness of various components. Here's an overview of the design and analysis process:

1. Problem Analysis:

Clearly define the problem scope, objectives, and constraints.
Analyze the data sources, including historical crop data, weather data, market data, and other relevant sources.
Identify key stakeholders and their requirements.

2. Data Collection and Preparation:

Collect and preprocess the data, including cleaning, normalization, and feature engineering.
Ensure data quality and address missing or erroneous data.
Determine data storage requirements and establish data pipelines.

3. Machine Learning Model Selection:

Choose appropriate machine learning algorithms based on the nature of the problem, data characteristics, and objectives.
Evaluate the suitability of deep learning models (e.g., neural networks) for tasks like image analysis or time series forecasting.

4. Infrastructure Design:

Design the hardware and software infrastructure, considering scalability, redundancy, and performance requirements.
Select the appropriate cloud or on-premises hosting environment.
Develop a plan for data backup, disaster recovery, and system monitoring.

5. Real-time Data Integration:

Design real-time data integration pipelines to ingest data from sources such as weather APIs, market feeds, and sensors.
Consider the use of stream processing frameworks for handling real-time data.

6. User Interface Design:

Design user interfaces for data visualization, reporting, and decision support.
Ensure that the user interface is intuitive and provides easy access to system insights and recommendations.

7. Security and Compliance:

Implement security measures such as data encryption, authentication, and access control.
Ensure compliance with relevant data protection and privacy regulations (e.g., GDPR).

Throughout the design and analysis process, consider involving domain experts, data scientists, software engineers, and stakeholders to ensure that the system aligns with the needs of the agriculture industry and effectively addresses the challenges of forecasting demand and supply information for major crops.

CHAPTER 7

IMPLEMENTATION

7. IMPLEMENTATION

Implementation is the stage where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that it will work efficiently and effectively.

The system can be implemented only after thorough testing is done and if it is found to work according to the specification. It involves careful planning, investigation of the current system and it constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods a part from planning.

Two major tasks of preparing the implementation are education and training of the users and testing of the system. The more complex the system being implemented, the more involved will be the system analysis and design effort required just for implementation.

The implementation phase comprises of several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed. For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

TESTING

The testing phase is an important part of software development. It is the Information zed system will help in automate process of finding errors and missing operations and also a complete verification to determine whether the objectives are met and the user requirements are satisfied. Software testing is carried out in three steps:

1. The first includes unit testing, where in each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether theobjectives have been met. Errors are noted down and corrected immediately.
2. Unit testing is the important and major part of the project. So errors are rectified easily in particular module and program clarity is increased. In this project entire system is dividedinto several modules and is developed individually. So unit testing is conducted to individual modules.
3. The second step includes Integration testing. It need not be the case, the software whose modules when run individually and showing perfect results, will also show perfect results when run as a whole.

CHAPTER 8

SNAPSHOTS

8. SNAPSHOTS

```
onion_data=pd.read_csv('C:/Users/Ullas/Crop Prediction/Potato.csv')
onion_data.info()
onion_data.head()
```

Python

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1827 entries, 0 to 1826
Data columns (total 10 columns):
 #   Column        Non-Null Count  Dtype 
---  ------        --------------  ----- 
 0   Sl no.        1827 non-null   int64 
 1   DistrictName  1827 non-null   object
 2   MarketName    1827 non-null   object
 3   Commodity     1827 non-null   object
 4   Variety       1827 non-null   object
 5   Grade         1827 non-null   object
 6   min_price     1827 non-null   int64 
 7   max_price     1827 non-null   int64 
 8   modal_price   1827 non-null   int64 
 9   PriceDate     1827 non-null   object
dtypes: int64(4), object(6)
memory usage: 142.9+ KB
```

	Sl no.	DistrictName	MarketName	Commodity	Variety	Grade	min_price	max_price	modal_price	PriceDate
0	1	Agra	Achnera	Potato	Desi	FAQ	900	1000	950	30-Dec-20
1	2	Agra	Achnera	Potato	Desi	FAQ	800	900	850	29-Dec-20
2	3	Agra	Achnera	Potato	Desi	FAQ	850	950	900	28-Dec-20
3	4	Agra	Achnera	Potato	Desi	FAQ	800	900	850	27-Dec-20
4	5	Agra	Achnera	Potato	Desi	FAQ	800	900	850	26-Dec-20

+ Code + Markdown

```
min_price=list(onion_data.min_price)
max_price =list(onion_data.max_price)
model_price =list(onion_data.modal_price)
arr =[list(onion_data.MarketName),list(onion_data.DistrictName), list(onion_data.Variety)]
index =pd.MultiIndex.from_arrays(arr, names=('MarketName','DistrictName','Variety'))
onion_df = pd.DataFrame({'minimum_price':min_price,'maximum_price':max_price,'modal_price':model_price}, index=index)
onion_df.head()
```

Python

			minimum_price	maximum_price	modal_price
MarketName	DistrictName	Variety			
Achnera	Agra	Desi	900	1000	950
		Desi	800	900	850
		Desi	850	950	900
		Desi	800	900	850
		Desi	800	900	850

```
onion_df2=onion_df.groupby(level=2,sort=False).mean().reset_index()
onion_df2.head()
```

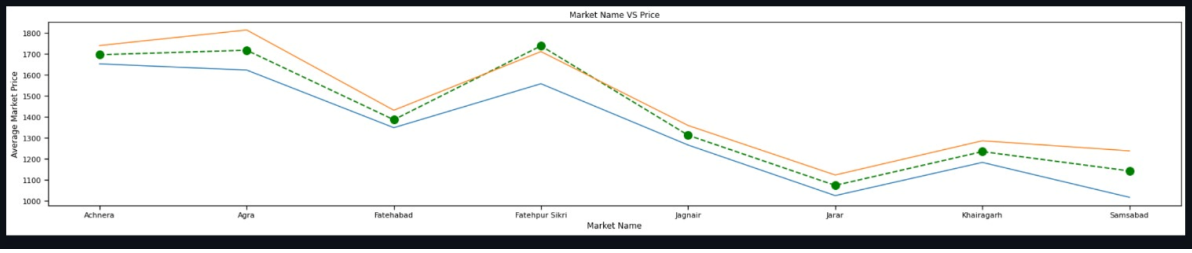
Python

	Variety	minimum_price	maximum_price	modal_price
0	Desi	1404.192251	1516.403509	1459.356725
1	Local	1358.427948	1536.681223	1518.788210
2	Other	1550.000000	1650.000000	1600.000000

```python
onion_df3=onion_df.groupby(level=0).mean().reset_index()

plt.figure(figsize=(30,5))
sns.set_context("notebook", font_scale=1)
plt.plot(onion_df3['MarketName'],onion_df3['modal_price'],color='green', marker='o', linestyle='dashed',linewidth=2, markersize=12)
plt.plot(onion_df3['MarketName'],onion_df3['minimum_price'])
plt.plot(onion_df3['MarketName'],onion_df3['maximum_price'])
plt.title("Market Name VS Price")
plt.xlabel("Market Name")
plt.ylabel("Average Market Price")
plt.show()
```

Python




```python
onion_df3.sort_values('modal_price',ascending=False)[0:5]
```

Python

	MarketName	minimum_price	maximum_price	modal_price
3	Fatehpur Sikri	1556.931034	1710.172414	1737.344828
1	Agra	1622.122807	1813.017544	1716.614035
0	Achnera	1651.633987	1738.921569	1695.212418
2	Fatehabad	1347.596899	1430.658915	1385.387597
4	Jagnair	1265.026738	1358.181818	1311.711230

```python
onion_df3.sort_values('modal_price')[0:5]
```

Python

	MarketName	minimum_price	maximum_price	modal_price
5	Jarar	1023.904110	1122.123288	1072.671233
7	Samsabad	1015.773810	1237.202381	1141.517857
6	Khairagarh	1182.085561	1285.294118	1233.689840
4	Jagnair	1265.026738	1358.181818	1311.711230
2	Fatehabad	1347.596899	1430.658915	1385.387597

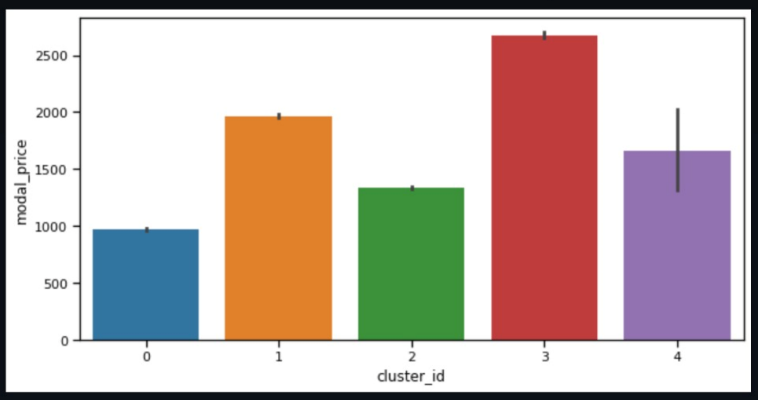
```python
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()
scaled_arr3 = scaler.fit_transform(onion_data[['modal_price','min_price','max_price']])

clusters = KMeans(5,random_state=42)# we've got k=5 using elbow plot
kmeans_model= clusters.fit(scaled_arr3)
onion_data['cluster_id']= kmeans_model.labels_
onion_data1 = onion_data.sort_values(['modal_price','cluster_id'])
onion_data1.head()
```

Python

	Sl no.	DistrictName	MarketName	Commodity	Variety	Grade	min_price	max_price	modal_price	PriceDate	cluster_id
764	765	Agra	Fatehabad	Potato	Desi	FAQ	500	600	550	11-May-20	0
816	817	Agra	Fatehabad	Potato	Desi	FAQ	500	600	550	13-Feb-20	0
1444	1445	Agra	Jarar	Potato	Desi	FAQ	600	700	650	03-Feb-20	0
1445	1446	Agra	Jarar	Potato	Desi	FAQ	600	700	650	02-Feb-20	0
1432	1433	Agra	Jarar	Potato	Desi	FAQ	640	720	680	17-Feb-20	0




```python
plt.figure(figsize=(10,5))
sns.set_context("notebook", font_scale=1)
sns.barplot(onion_data1['cluster_id'],onion_data1['modal_price'])
plt.show()
```

Python

CHAPTER 9

CONCLUTION

9. CONCLUTION

The package was designed in such a way that future modifications can be done easily. The following conclusions can be deduced from the development of the project:

- Automation of the entire system improves the efficiency
- It provides a friendly graphical user interface which proves to be better when compared to the existing system.
- It gives appropriate access to the authorized users depending on their permissions.
- It effectively overcomes the delay in communications.
- Updating of information becomes so easier
- System security, data security and reliability are the striking features.
- The System has adequate scope for modification in future if it is necessary.

10. REFERENCE

1.Arjun KM. Indian agriculture-status, importance and role in Indian economy. International Journal of Agriculture and Food Science Technology, 2020: 4(4): 343–346.

2. Choudhary B, Gaur K. The development and regulation of Bt brinjal in India (Eggplant/Aubergine). International Service for the Acquisition of Agri-biotech Applications. 2018.

3. Fafchamps M, Minten B. Impact of SMS-based agricultural information on Indian farmers. The World Bank Economic Review, 2022; 26(3): 383–414.

4. Xiong T, Li C, Bao Y. Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China. Neurocomputing, 2019; 275, 2831–2844.

5. Sankaran S. Demand forecasting of fresh vegetable product by seasonal ARIMA model. International Journal of Operational Research 20, 2019; (3): 315–330.

6. Paul RK, Panwar S, Sarkar SK, Kumar A, Singh KN, Farooqi S, et al. Modelling and forecasting of meat exports from India. Agricultural Economics Research Review 2019; 26 (2): 249–256.

7. Flach P. Machine Learning: The Art and Science of Algorithms that Make Sense of Data. 2022: Cambridge University Press, Cambridge.

8. Werbos PJ. Generalization of backpropagation with application to a recurrent gas market model, Neural Network. 2020; 1(4): 339–356.

9. Bokde ND, Yaseen ZM, Gorm BA. ForecastTB—An R Package as a Test-Bench for Time Series Forecasting—Application of Wind Speed and Solar Radiation Modeling. Energies, 2020; 13: 10: 2578.

10. Hansen PR, Lunde A, Nason JM. The model confidence set. Econometrica, 2021; 79(2): 453–497.