

Smart Cost

**Hybrid ARIMA Prophet LSTM Volatility Forecasting with
Agentic News Fusion for Commodity Price Prediction in
Resource Constrained SMEs**

R26-IT-026

Project Proposal Report
Jayasingha D.D.M – IT22162250

B.Sc. (Hons) Degree in Information Technology Specialized in
Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

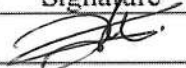
March 2026

DECLARATION

I declare that this is my own work, and this proposal does not knowingly incorporate any material previously submitted for a degree or diploma at any other university or higher education institution, nor does it contain any material that has been previously published or written by another person apart from where proper acknowledgment is given in the text.

Name	Student ID	Signature
Jayasingha D.D.M	IT22162250	

The above candidate is conducting research for their undergraduate Dissertation under my supervision.

Name	Date	Signature
Mr. Deemantha Siriwardana	03/12/2026	

ABSTRACT

Small and micro-sized businesses (micro-SMEs), especially bakeries and small food manufacturing companies, rely significantly on raw ingredients such as wheat, sugar, butter, and cooking oil, the prices of which are subject to regular fluctuations because of supply chain instability, seasonal demand, and economic variables. However, the majority of microbusinesses control ingredient costs through manual observation or short-term experience, which frequently results in lower profit margins and delayed responses to price fluctuations. While big organizations use sophisticated ERP systems and predictive analytics tools to predict changes in supply chain prices, small businesses frequently cannot afford or access these technologies. The identified research need is the absence of lightweight, reasonably priced predictive analytics tools that can predict changes in ingredient prices using the scant historical data that micro-SMEs usually have access to. Current inventory or financial management systems are primarily concerned with documenting past transactions rather than offering projections of future cost fluctuations. This study suggests a predictive analytics module that makes use of time-series forecasting methods to discover supply chain volatility patterns and estimate future trends in ingredient prices. In order to break down patterns, seasonality, and market volatility, the suggested method would examine historical commodity price statistics and use statistical forecasting models like Facebook Prophet and Auto Regressive Integrated Moving Average (ARIMA). Additionally, the system will use confidence interval analysis to produce early alerts regarding possible price hikes. The forecasting module will be implemented using Python-based technologies including Pandas, NumPy, Statsmodels, and Prophet, and will expose prediction services through a FastAPI backend for integration with the mobile decision-support system. The expected outcome of this research is a lightweight forecasting engine capable of generating reliable ingredient price predictions and volatility alerts. The system will assist small business owners in making proactive purchasing decisions, optimizing inventory planning, and maintaining stable profit margins. From a broader perspective, the research contributes to the digital transformation of micro-SMEs by democratizing predictive analytics tools that were previously accessible only to large enterprises.

Keywords : Supply Chain Forecasting, Time-Series Prediction, Micro-SME Analytics, ARIMA, Facebook Prophet.

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere gratitude to everyone who has contributed to the development of this project proposal for the Smart Cost : Smart Supply Chain Cost Analytics and Profit Prediction System for Micro SMEs (Predictive Supply Chain Volatility Forecasting).

First and foremost, I would like to thank my supervisor, Mr. Deemantha Siriwardana, for his valuable guidance, encouragement, and continuous support throughout the preparation of this research proposal. His expertise and insightful advice have been extremely helpful in understanding the research problem and refining the predictive analytics approach proposed in this study.

Next, I would also like to thank my cosupervisor, Miss Ayesha Wijesooriya, for her continuous support, constructive feedback, and encouragement during the development of this research proposal. Her guidance and academic insights have been highly valuable in improving the quality and direction of this research work.

Then, I would like to thank my colleagues and team members for their cooperation, support, and valuable discussions during the early stages of this research project. Their willingness to share ideas and collaborate has helped strengthen the research concept and improve the overall project direction.

Lastly, I would like to extend my sincere appreciation to the Faculty of Computing and the Comprehensive Design and Analysis Project (CDAP) team at the Sri Lanka Institute of Information Technology (SLIIT) for providing me with the opportunity to conduct this research. Their support, academic guidance, and the resources provided have played an important role in enabling the successful preparation of this project proposal.

TABLE OF CONTENT

DECLARATION	2
ABSTRACT	3
ACKNOWLEDGEMENT	4
LIST OF TABLES	7
LIST OF FIGURES	8
LIST OF ABBREVIATIONS	9
LIST OF APPENDICES.....	10
1. INTRODUCTION.....	11
1.1 Background and Context	11
1.2 Literature Review	12
1.3 Research Gap	13
2. OBJECTIVES	15
2.1 Main Objective	15
2.2 Specific Objectives	15
3. METHODOLOGY	16
3.1 Research Approach	16
3.2 Technical Methods and Algorithms	16
3.3 Tools and Platforms	19
3.4 Data Collection and Ethical Considerations	20
3.5 Validation Strategy and Performance Metrics.....	21
4. HIGH-LEVEL SYSTEM ARCHITECTURE	23
4.1 System Overview	24
4.2 System Component Diagram	25
4.3 Component Interactions	25
4.4 Data Flow Explanation	25
4.5 Scalability and Security Considerations	27
5. USER REQUIREMENTS	28
5.1 Stakeholder Identification	28
5.2 Functional Requirements	28
5.3 Non-Functional Requirements	29
6. COMMERCIALIZATION PLAN	31
6.1 Target Market and Customer Persona	31
6.2 Value Proposition	33
6.3 Revenue Model	33
6.4 Cost Estimation.....	33
6.5 Pricing Strategy.....	34
6.6 Competitive Advantage.....	34
6.7 Intellectual Property Considerations.....	35
7. BUDGET AND JUSTIFICATION	36
7.1 Hardware Costs.....	36

7.2 Software Licenses	36
7.3 Cloud Services	37
7.4 Human Effort Cost.....	37
8. WORK BREAKDOWN STRUCTURE (WBS)	40
9. GANTT CHART	41
10. REFERENCES.....	42
11. APPENDICES	42
A – Survey Instruments.....	43
B – System Diagrams	44
C – UI Mockups.....	44
D – Risk Analysis	45
E – Ethical Clearance Documents.....	46

LIST OF TABLES

Table 1: Comparison of Time-Series Forecasting Models (ARIMA vs Prophet).....	19
Table 2: Summary of Historical Commodity Price Datasets Used for Forecasting....	20
Table 3: Data Preprocessing Steps for Ingredient Price Dataset.....	21
Table 4: Evaluation Metrics Used for Model Validation.....	22
Table 5: Hardware Costs.....	36
Table 6: Software Licenses.....	37
Table 7: Cloud Services.....	38
Table 8: Human Effort Cost.....	38
Table 9: Total Estimated Project Budget.....	39
Table 10: Work Breakdown Structure for Predictive Analytics Component.....	40

LIST OF FIGURES

Figure 1: Data Collection and Preprocessing Workflow.....	18
Figure 2: Time-Series Forecasting Process Using ARIMA Model.....	18
Figure 3: Supply Chain Volatility Detection using Confidence Intervals.....	18
Figure 4: Ingredient Price Trend Analysis Example.....	22
Figure 5: Overall Predictive Supply Chain Forecasting System Architecture.....	23
Figure 6: Integration of Forecasting Module with Mobile Decision Support System...	25
Figure 7: Forecasted Price Trends Visualization for Bakery Ingredients.....	26
Figure 8: Proposed Predictive Analytics Data Flow.....	26
Figure 9: WBS Diagram.....	40
Figure 10: Project Gantt Chart Timeline.....	41
Figure 11: System Diagram 01.....	44
Figure 12: System Diagrams 02.....	44

LIST OF ABBREVIATIONS

ARIMA	Auto Regressive Integrated Moving Average
API	Application Programming Interface
AI	Artificial Intelligence
ERP	Enterprise Resource Planning
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
SME	Small and Medium-Sized Enterprise
Micro-SME	Micro Small and Medium-Sized Enterprise
NLP	Natural Language Processing
OCR	Optical Character Recognition
WBS	Work Breakdown Structure
SDG	Sustainable Development Goals

LIST OF APPENDICES

Appendix A – Survey Instruments.....	43
Appendix B – Detailed System Diagrams.....	44
Appendix C – Preliminary UI Mockups.....	44
Appendix D – Risk Analysis.....	45
Appendix E – Ethical Clearance Document.....	46

1. INTRODUCTION

1.1 Background and Context

Small and micro-sized enterprises (micro-SMEs) play a significant role in global and local economies, particularly in sectors such as food production, retail, and small-scale manufacturing. Among these businesses, bakeries and small food production companies heavily depend on raw materials such as wheat flour, sugar, butter, eggs, and cooking oil to produce their daily products. The prices of these essential ingredients are highly volatile due to multiple factors, including supply chain disruptions, seasonal demand variations, global market instability, and economic inflation. As a result, frequent price fluctuations directly impact the operational costs and profitability of micro-scale businesses.

Despite the increasing volatility of raw material prices, many micro-SMEs still rely on traditional methods such as manual monitoring, past experience, or informal market observation to manage ingredient costs. These approaches often lack accuracy and predictive capability, causing delayed responses to price changes. Consequently, small business owners may purchase ingredients at unfavorable prices or fail to adjust product pricing in time, leading to reduced profit margins or financial instability.

Globally, small and medium enterprises account for approximately 90% of all businesses and contribute more than 50% of global employment, according to the World Bank. In developing economies, micro-SMEs contribute significantly to national economic growth and employment opportunities. In Sri Lanka, the food and bakery industry represents a substantial portion of the micro-enterprise sector, with thousands of small bakeries operating across the country. However, these businesses often operate with limited technological support and lack access to advanced financial forecasting tools that could assist in managing fluctuating ingredient costs.

Large enterprises typically utilize advanced Enterprise Resource Planning (ERP) systems and predictive analytics platforms to forecast supply chain trends and manage procurement strategies. These systems use sophisticated data analytics and machine learning techniques to anticipate price fluctuations and optimize purchasing decisions. However, such technologies are often expensive, technically complex, and require specialized infrastructure, making them inaccessible to most micro-SMEs.

As a result, there is a growing need for lightweight and affordable predictive analytics solutions specifically designed for small businesses. The proposed research addresses this challenge by developing a predictive supply chain volatility forecasting module capable of analyzing historical ingredient price data and identifying future trends. By applying time-series forecasting models such as Auto Regressive Integrated Moving Average (ARIMA) and Facebook Prophet, the system will analyze historical commodity price patterns and generate predictions about future price fluctuations. The forecasting module will also incorporate confidence interval analysis to identify potential supply chain volatility and generate early warning alerts for business owners.

The proposed solution will be implemented using Python-based technologies, including Pandas, NumPy, Statsmodels, and Prophet, and will expose prediction services through a FastAPI backend for integration with a mobile decision-support system. This system will enable micro-SME owners to access predictive insights about ingredient price trends, allowing them to make proactive purchasing decisions and maintain stable profit margins.

This research also aligns with the United Nations Sustainable Development Goals (SDGs), particularly SDG 8: Decent Work and Economic Growth, which promotes sustainable economic development and productivity improvement for small enterprises. By providing predictive analytics tools to micro-businesses, the proposed system supports digital transformation and strengthens the resilience of small-scale enterprises in volatile economic environments. From an academic perspective, this research falls within the Technology Integration and Management research cluster, as it integrates predictive

analytics, time-series forecasting, and decision-support systems to solve a real-world supply chain management problem. By combining these technologies into a lightweight and accessible platform, the research contributes to the development of intelligent decision-support tools tailored for micro-SME environments.

1.2 Literature Review

The increasing volatility of global supply chains and fluctuating commodity prices have attracted significant research attention in recent years. Businesses that depend heavily on raw materials often experience financial uncertainty due to sudden price changes in essential commodities such as agricultural products, energy resources, and industrial materials. As a result, researchers and practitioners have explored various predictive analytics techniques to forecast price movements and assist organizations in making informed procurement and inventory management decisions. Time-series forecasting models, machine learning techniques, and statistical analysis methods have been widely applied in supply chain analytics to address these challenges and support data-driven decision making.

Several studies have explored the application of traditional statistical models for commodity price forecasting. One of the most widely used approaches is the Autoregressive Integrated Moving Average (ARIMA) model introduced by Box and Jenkins [1]. ARIMA models analyze historical time-series data by capturing temporal relationships between past observations and future values. This technique has been extensively applied in economic forecasting, financial market prediction, and commodity price analysis due to its ability to model trends and temporal dependencies within structured datasets. Research conducted by Zhang et al. demonstrated that ARIMA-based forecasting models can effectively predict agricultural commodity prices when sufficient historical data is available [2]. Their findings show that statistical time-series models remain reliable tools for forecasting commodity price trends.

In recent years, new forecasting frameworks have been developed to improve prediction accuracy and handle complex time-series patterns. One notable example is Facebook Prophet, a forecasting framework specifically designed for business time-series analysis [3]. Prophet decomposes time-series data into several components, including trend, seasonality, and holiday effects. This approach allows the model to generate accurate forecasts even when the dataset contains irregular patterns or missing observations. Taylor and Letham demonstrated that Prophet performs effectively in forecasting retail demand, supply chain metrics, and business operational data with relatively low computational complexity [3]. As a result, Prophet has become widely adopted in business forecasting applications.

Machine learning approaches have also been explored for supply chain forecasting tasks. Neural network models such as Long Short-Term Memory (LSTM) networks have shown promising results in predicting complex time-series patterns [4]. LSTM models are capable of capturing long-term dependencies in sequential data, making them suitable for applications such as stock market prediction, energy demand forecasting, and commodity price analysis. However, these models typically require large volumes of training data and high computational resources. For small-scale enterprises that operate with limited historical data and computational capacity, the implementation of deep learning models may not be practical.

Several researchers have also examined broader forecasting methodologies and their applications in business environments. Hyndman and Athanasopoulos emphasized that selecting an appropriate forecasting model depends on the characteristics of the dataset and the forecasting objective [9]. Their work highlights that traditional statistical models remain effective for many forecasting applications, particularly when datasets are relatively small and structured. In addition, Hyndman and Khandakar proposed automated time-series forecasting techniques that combine statistical modeling with algorithmic model selection to improve forecasting performance [10]. These studies demonstrate that

statistical forecasting techniques continue to play an important role in practical predictive analytics applications.

In addition to academic research, several commercial supply chain analytics platforms have been developed to support forecasting and procurement planning. Enterprise Resource Planning (ERP) systems such as SAP Integrated Business Planning, Oracle Supply Chain Management, and IBM Supply Chain Analytics provide advanced forecasting capabilities for large enterprises [5]. These platforms integrate historical data, forecasting algorithms, and procurement planning tools to optimize supply chain operations and improve business performance. However, the implementation of such systems typically requires substantial financial investment, complex system integration, and specialized technical expertise. Consequently, these enterprise-level solutions are often inaccessible to micro and small-scale enterprises.

Similarly, several cloud-based inventory management systems provide basic forecasting functionality. Platforms such as NetSuite, TradeGecko, and Zoho Inventory offer features that support stock monitoring and demand forecasting [6]. While these tools provide valuable support for inventory management, their forecasting capabilities primarily focus on predicting product demand rather than forecasting raw material price volatility. As a result, they do not fully address the challenges faced by businesses that depend on fluctuating commodity prices for their production processes.

Recent studies have also highlighted the challenges faced by small and medium-sized enterprises in adopting advanced data analytics technologies. According to research conducted by Choi et al., many small businesses encounter barriers such as high technology costs, limited technical expertise, and insufficient data infrastructure when attempting to implement predictive analytics systems [7]. Similarly, Wang and Zhang emphasized that predictive analytics tools designed for small enterprises must operate effectively with limited datasets while maintaining reasonable levels of forecasting accuracy [8]. These findings indicate that lightweight and cost-effective predictive solutions are necessary to support decision-making in small business environments.

Despite the growing availability of forecasting techniques and enterprise analytics platforms, there remains a significant gap in predictive analytics solutions specifically designed for micro and small-scale enterprises (micro-SMEs) operating in sectors such as food production and retail. Most existing forecasting systems either rely on large datasets, require advanced technical infrastructure, or involve expensive enterprise software licenses. Consequently, many micro-business owners still rely on manual observation or past experience when managing ingredient costs, which often results in delayed responses to price fluctuations.

Therefore, this research proposes the development of a lightweight predictive supply chain volatility forecasting module tailored for micro-SME environments. By utilizing statistical time-series forecasting models such as ARIMA and Facebook Prophet, the proposed system aims to generate accurate ingredient price predictions using relatively small historical datasets. This approach addresses the limitations of existing forecasting systems while making predictive analytics more accessible and practical for small business environments.

1.3 Research Gap

Despite the growing research in supply chain analytics and commodity price forecasting, there remains a significant gap in predictive analytics solutions specifically designed for micro and small-scale enterprises (micro-SMEs). Existing studies have shown that time-series forecasting models such as Autoregressive Integrated Moving Average (ARIMA) and other advanced analytical techniques can effectively predict commodity price trends when large historical datasets and advanced computational resources are available. However, most of these studies focus on large-scale industries, agricultural commodity markets, or financial trading environments where extensive datasets and sophisticated analytical infrastructures are accessible.

In contrast, micro-SMEs such as small bakeries and food production businesses operate under limited technological and financial resources. These businesses typically maintain only basic purchase records and lack the technical infrastructure required to implement complex predictive analytics systems. As a result, many small business owners rely on manual observations or past experience when making procurement decisions. This reactive approach often leads to delayed responses to ingredient price fluctuations and reduced profit margins.

Although enterprise-level forecasting platforms such as SAP Supply Chain Management and Oracle Supply Chain Analytics provide advanced predictive capabilities, their high implementation costs and technical complexity make them impractical for micro-SMEs. Similarly, existing cloud-based inventory management systems such as Zoho Inventory primarily focus on stock tracking and sales demand forecasting rather than predicting raw material price volatility.

Another challenge is that many modern forecasting techniques rely on large volumes of structured data, which are rarely available in small business environments. Consequently, there is a need for lightweight forecasting solutions that can generate reliable predictions using limited historical datasets.

Therefore, this research addresses the identified gap by developing a predictive supply chain volatility forecasting module tailored for micro-SMEs. The proposed system utilizes statistical time-series forecasting models, including ARIMA and Prophet, combined with confidence interval analysis to detect potential price fluctuations and generate early warning alerts. This approach aims to provide an affordable and practical predictive analytics solution that enables micro-SMEs to anticipate ingredient price volatility and make informed procurement decisions.

2. OBJECTIVE

2.1 Main Objective

To design and implement a predictive supply chain volatility forecasting module that analyzes historical ingredient price data and predicts future price trends using time-series models such as ARIMA and Prophet, enabling micro and small-scale enterprises (micro-SMEs) to make proactive purchasing decisions and manage raw material price fluctuations.

2.2 Specific Objectives

Objective 1 – Collect and Prepare Historical Ingredient Price Data

To collect and preprocess historical commodity price datasets for key bakery ingredients such as flour, sugar, butter, and cooking oil from publicly available sources, and transform the data into a structured time-series format suitable for forecasting analysis.

Objective 2 – Develop Time-Series Forecasting Models

To implement and evaluate statistical forecasting models, including ARIMA and Facebook Prophet, for analyzing historical ingredient price trends and generating accurate future price predictions.

Objective 3 – Detect Supply Chain Volatility Patterns

To analyze historical price fluctuations and detect supply chain volatility patterns by decomposing time-series data into trend, seasonality, and irregular components and identifying potential price instability.

Objective 4 – Evaluate Forecasting Model Performance

To assess the accuracy and reliability of the forecasting models using evaluation metrics such as Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) to ensure reliable predictions for micro-SMEs.

3. METHODOLOGY

3.1 Research Approach

This research follows a System Development and Experimental Evaluation methodology. The system development component focuses on designing and implementing a predictive supply chain volatility forecasting module, while experimental evaluation is used to compare forecasting models and measure prediction accuracy using quantitative performance metrics. The Design Science Research approach focuses on designing and developing an artifact that addresses a real-world problem, while experimental evaluation is used to assess the effectiveness of the proposed solution using quantitative performance metrics.

In this research, the artifact is a predictive supply chain volatility forecasting module capable of identifying patterns in historical ingredient price data and predicting future price trends. The forecasting module is designed to assist micro and small-scale enterprises (micro-SMEs), such as bakeries and small food production businesses, in anticipating raw material price fluctuations and making proactive purchasing decisions.

The research begins with identifying the real-world problem of fluctuating ingredient prices affecting micro-SMEs. Many small businesses currently rely on manual monitoring or past experience to manage ingredient costs, which often leads to delayed responses to price changes and reduced profit margins. To address this challenge, the proposed system is designed to provide predictive insights using time-series forecasting techniques.

The research process consists of several stages. First, historical ingredient price datasets are collected from publicly available sources such as government statistical portals and commodity price databases. These datasets are then preprocessed to ensure data quality and converted into structured time-series format suitable for forecasting analysis.

Next, statistical forecasting models are applied to analyze historical patterns and generate future price predictions. The two primary forecasting models used in this research are Auto Regressive Integrated Moving Average (ARIMA) and the Prophet forecasting model. These models are trained using historical price datasets and evaluated using quantitative performance metrics.

To evaluate forecasting accuracy, the dataset will be divided into training and testing subsets. The models will be trained using historical training data and then evaluated on the testing dataset to measure prediction accuracy.

The performance of the ARIMA and Prophet models will be compared using standard evaluation metrics such as Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). This experimental evaluation helps determine the most suitable forecasting approach for predicting ingredient price volatility in micro-SME environments.

Finally, the forecasting module will be deployed as a backend prediction service using FastAPI, enabling the forecasting results to be accessed by a mobile decision-support system. Through this integration, micro-SME owners will be able to view ingredient price predictions and receive early warning alerts about potential supply chain volatility.

3.2 Technical Methods and Algorithms

The forecasting module utilizes time-series analysis techniques to identify patterns in historical ingredient price data and generate predictions about future price movements. Two widely used

forecasting models are implemented in this research: the Auto Regressive Integrated Moving Average (ARIMA) model and the Prophet forecasting model.

The ARIMA model is a statistical forecasting method that analyzes historical time-series data by capturing relationships between past observations and future values. ARIMA models are represented as $ARIMA(p, d, q)$, where:

- p represents the number of autoregressive terms
- d represents the order of differencing used to make the time series stationary
- q represents the number of moving average terms

These parameters allow the ARIMA model to capture trends and temporal dependencies within historical price data. ARIMA has been widely used in economic forecasting, financial analysis, and commodity price prediction due to its effectiveness in modeling structured time-series data.

In addition to ARIMA, the Prophet forecasting model will be implemented to improve forecasting flexibility and accuracy. Prophet is designed specifically for business time-series forecasting and can effectively handle irregular data patterns, missing values, and seasonal variations. The model decomposes time-series data into three main components:

- Trend component – long-term growth or decline in the data
- Seasonality component – repeating patterns over time
- Holiday or event component – effects caused by specific events or irregular factors

By decomposing the data into these components, the Prophet model can generate reliable forecasts even when the data contains noise or irregular patterns.

The forecasting process begins with data preprocessing, which includes handling missing values, smoothing noisy observations, and converting raw datasets into structured time-series format. Preprocessing ensures that the dataset is suitable for training forecasting models.

Once preprocessing is completed, both ARIMA and Prophet models are trained using historical ingredient price datasets. The trained models generate future price predictions along with confidence intervals, which represent the uncertainty range of the forecasted values.

Confidence interval analysis is also used to detect potential supply chain volatility events. If the predicted upper confidence bound exceeds the current market price beyond a predefined threshold, the system generates an early warning signal indicating possible price increases. These alerts allow business owners to take proactive procurement actions before ingredient prices rise significantly.

By combining statistical time-series forecasting techniques with confidence interval analysis, the proposed system can identify potential supply chain instability and provide predictive insights that support better purchasing and inventory management decisions for micro-SMEs

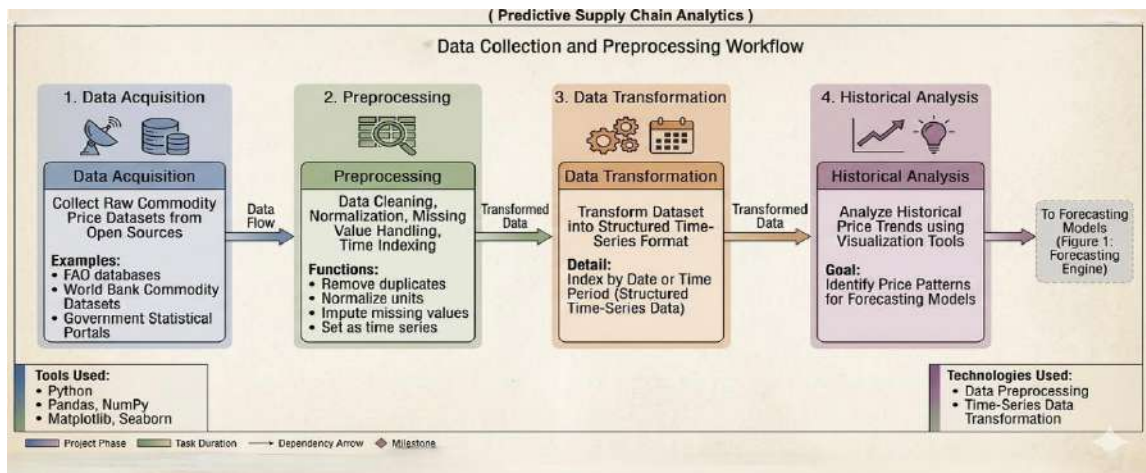


Figure 1

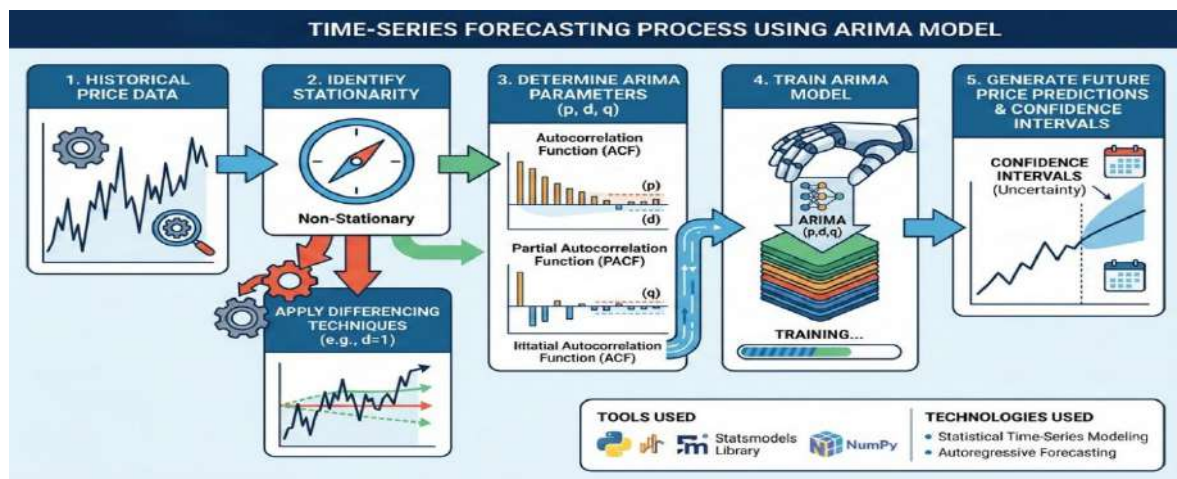


Figure 2

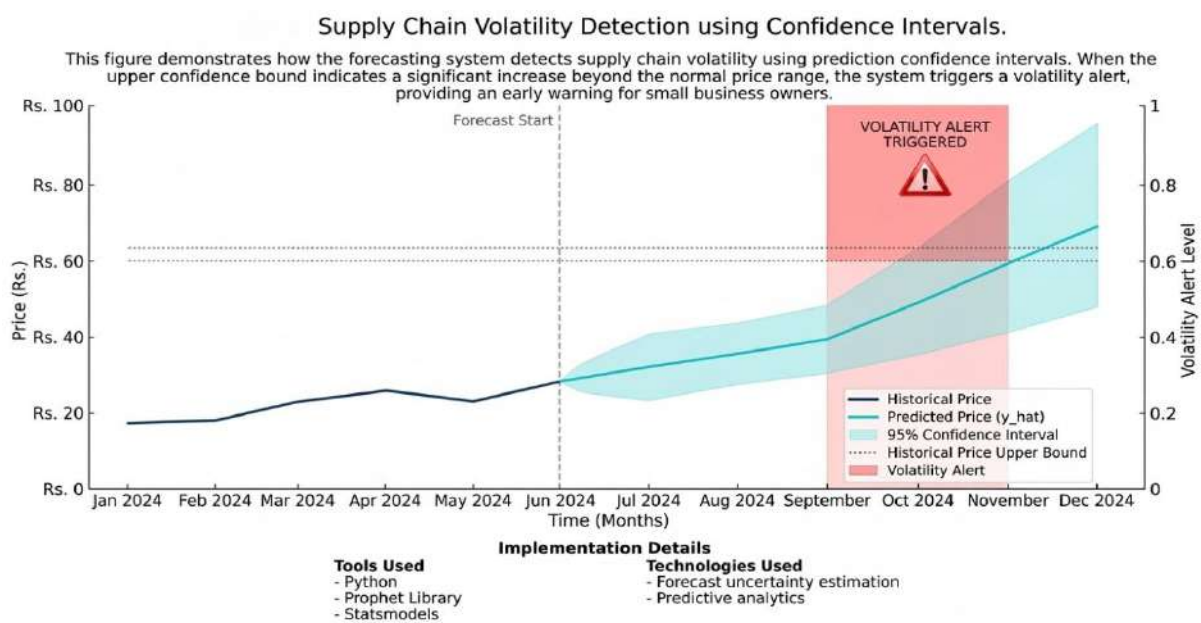


Figure 3

Table 1: Comparison of Time-Series Forecasting Models (ARIMA vs Prophet)

Feature	ARIMA Model	Facebook Prophet
Model Type	Statistical Time Series Model	Additive Time Series Model
Best For	Stationary data	Business time-series forecasting
Handles Seasonality	Limited	Built-in seasonal components
Handles Missing Data	Requires preprocessing	Handles missing data automatically
Data Requirement	Requires clean dataset	Works with irregular datasets
Complexity	Moderate	Low
Libraries Used	Statsmodels	Prophet
Programming Language	Python	Python
Output	Forecasted values with confidence intervals	Forecast trend + seasonality + uncertainty

3.3 Tools and Platforms

The proposed forecasting module will be developed using Python-based technologies, which provide powerful libraries for data analysis and machine learning. Python is widely used in predictive analytics and time-series forecasting due to its extensive ecosystem of data science libraries.

The primary libraries used in this research include Pandas and NumPy, which will be used for data manipulation and numerical analysis. Pandas provides powerful tools for handling time-series datasets, while NumPy supports mathematical operations required for statistical analysis.

The Statsmodels library will be used to implement the ARIMA forecasting model. This library provides statistical modeling capabilities that enable the development and evaluation of time-series forecasting algorithms.

The Facebook Prophet library will be used to develop the alternative forecasting model capable of handling seasonality and irregular trends within business datasets. Prophet simplifies the process of building forecasting models while maintaining high accuracy for time-series predictions.

To expose forecasting results to external systems, a lightweight backend service will be developed using FastAPI, a modern Python web framework designed for building high-performance APIs. The FastAPI service will allow the forecasting module to provide prediction results to a mobile decision-support system through API requests.

Visualization libraries such as Matplotlib and Seaborn will also be used to generate graphical representations of ingredient price trends and forecasting results, which can be integrated into the mobile system dashboard.

3.4 Data Collection and Ethical Considerations

The forecasting module requires historical ingredient price datasets to analyze supply chain trends and train the forecasting models. The required datasets will be obtained from publicly available sources such as government statistical portals, open data repositories, and online commodity price datasets.

Examples of data sources include national statistical reports and open datasets that provide historical price records for commonly used commodities such as wheat flour, sugar, cooking oil, and dairy products. These datasets typically include information such as price values, timestamps, and commodity categories.

Before analysis, the collected datasets will undergo preprocessing to ensure data quality. This includes removing duplicate entries, handling missing values, and normalizing measurement units to maintain consistency across datasets. Since the data used in this research originates from publicly available sources and does not contain sensitive personal information, the ethical risk associated with data usage is minimal. However, the research will still adhere to responsible data handling practices, ensuring that all datasets are used solely for academic research purposes. Proper acknowledgment will also be given to the original data providers.

Table 2: Summary of Historical Commodity Price Datasets Used for Forecasting

Dataset Name	Ingredient Type	Data Source	Data Format	Time Range	Attributes	Purpose
Global Wheat Price Dataset	Wheat Flour	FAO / World Bank Commodity Price Data	CSV / Excel	2015–2024	Date, Price per Kg, Market Index	Used to analyze flour price trends
Sugar Commodity Price Dataset	Sugar	World Bank Commodity Price Database	CSV	2015–2024	Date, Price per Kg, Market Location	Forecast sugar price volatility
Dairy Product Price Dataset	Butter / Milk	National Agriculture Statistics	CSV	2016–2024	Date, Price per Unit, Supplier Data	Predict butter price trends
Cooking Oil Market Dataset	Cooking Oil	Kaggle Commodity Dataset	CSV	2017–2024	Date, Price per Liter	Analyze oil price fluctuations
Local Market Price Dataset	Mixed Ingredients	Government Statistical Portal	CSV / JSON	2018–2024	Date, Ingredient Name, Price	Local market prediction training

Table 3: Data Preprocessing Steps for Ingredient Price Dataset

Step	Description	Tools / Libraries Used	Output
Data Collection	Import raw commodity price data	Python, Pandas	Raw dataset
Data Cleaning	Remove duplicates and invalid values	Pandas	Clean dataset
Missing Value Handling	Fill missing values using interpolation	Pandas / NumPy	Complete dataset
Data Normalization	Standardize measurement units	NumPy	Normalized dataset
Time Series Conversion	Convert data to time-indexed format	Pandas Datetime Index	Time-series dataset
Data Visualization	Analyze historical trends	Matplotlib / Seaborn	Trend graphs

3.5 Validation Strategy and Performance Metrics

To evaluate the effectiveness of the forecasting models, a validation strategy will be implemented using historical data splitting techniques. The dataset will be divided into two parts: a training dataset and a testing dataset. The forecasting models will be trained using the training dataset and then evaluated using the testing dataset to measure prediction accuracy.

Several performance metrics will be used to assess the quality of the forecasting models. One of the primary evaluation metrics used in this research is Mean Absolute Percentage Error (MAPE), which measures the average percentage difference between predicted values and actual values. A lower MAPE value indicates higher prediction accuracy.

Another important evaluation metric is Root Mean Square Error (RMSE), which measures the magnitude of prediction errors. RMSE is particularly useful for identifying large forecasting errors and evaluating overall model performance.

By comparing the performance of the ARIMA and Prophet models using these metrics, the research will determine the most suitable forecasting approach for predicting ingredient price fluctuations in micro-SME environments.

The validation results will provide evidence that the proposed forecasting module can generate reliable predictive insights that assist micro-business owners in anticipating supply chain price volatility and making informed purchasing decisions.

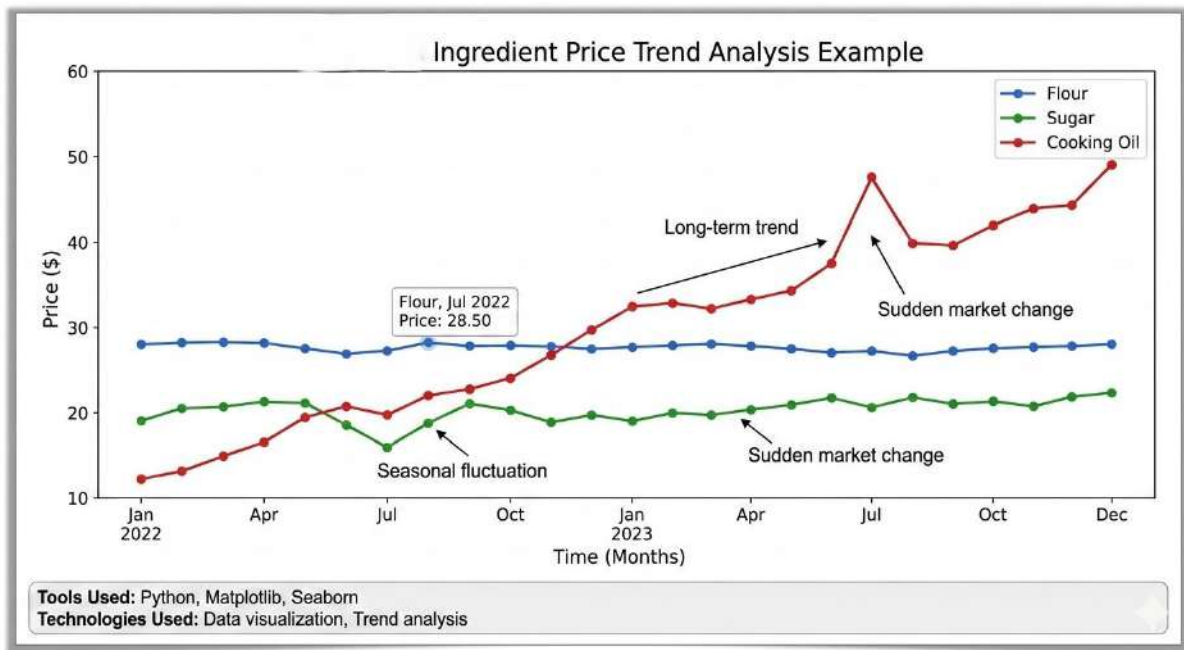


Figure 4

Table 4: Evaluation Metrics Used for Model Validation

Metric	Description	Formula	Purpose
Mean Absolute Percentage Error (MAPE)	Measures prediction accuracy as a percentage	$MAPE = (1/n) \sum$	Actual - Predicted
Root Mean Square Error (RMSE)	Measures magnitude of prediction errors	$RMSE = \sqrt{(\sum(\text{Predicted} - \text{Actual})^2 / n)}$	Identify large prediction errors
Mean Absolute Error (MAE)	Average of absolute prediction errors	$MAE = \sum$	Predicted - Actual
R-Squared Score	Measures goodness of fit	$R^2 = 1 - (SS_{\text{res}} / SS_{\text{tot}})$	Evaluate model performance

4. HIGH-LEVEL SYSTEM ARCHITECTURE

4.1 System Overview

The proposed system is designed to analyze historical ingredient price datasets and generate predictions about future supply chain price trends. The system architecture consists of several functional layers, including data acquisition, preprocessing, forecasting analysis, prediction service delivery, and visualization integration.

The system begins with the data collection layer, where historical commodity price datasets for ingredients such as flour, sugar, butter, and cooking oil are gathered from publicly available sources. These datasets are then processed by a data preprocessing module that cleans and structures the data into a time-series format suitable for forecasting analysis.

After preprocessing, the structured data is passed to the forecasting engine, which implements statistical forecasting models such as ARIMA and Facebook Prophet. These models analyze historical patterns and generate predictions of future ingredient price movements.

The predicted results, along with confidence intervals and volatility indicators, are then delivered through a prediction service API implemented using FastAPI. This API allows external applications, such as mobile decision-support systems, to retrieve forecasting results and display them to users through interactive dashboards or alerts.

The overall system architecture is designed to provide accurate predictive insights while remaining lightweight and accessible for small business environments.

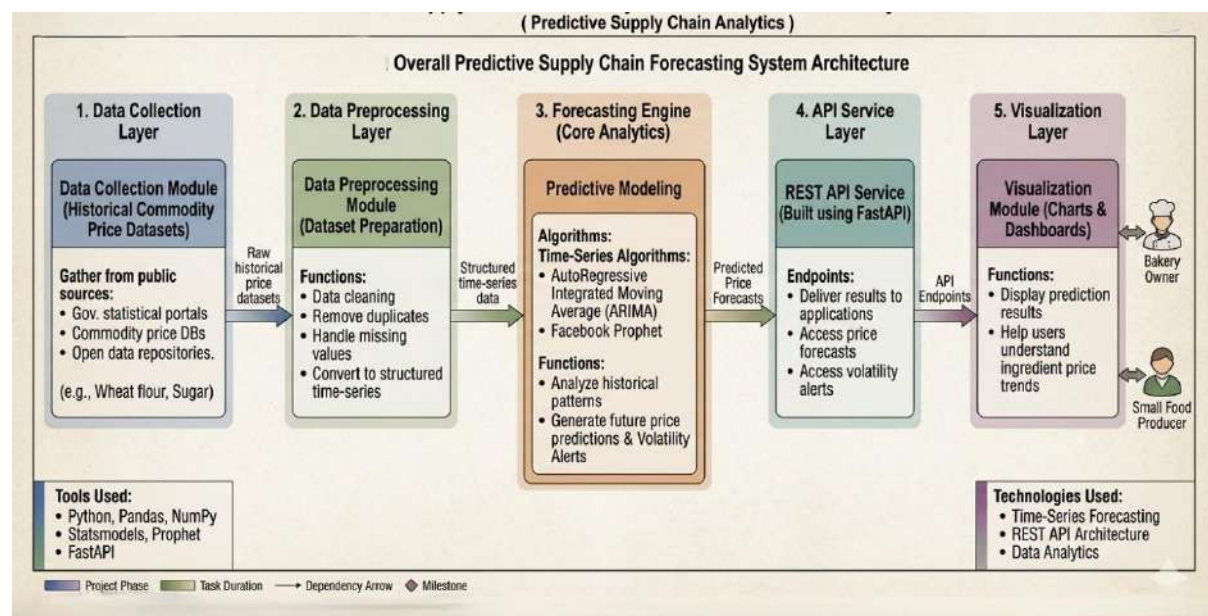


Figure 5

4.2 System Component Diagram

The high-level system architecture can be represented through several key components that interact to generate predictive analytics results. The major components of the system include:

1. Data Collection Module

This component gathers historical ingredient price datasets from public data repositories, government statistical portals, and commodity price databases. The collected data includes timestamps, ingredient names, and recorded market prices.

2. Data Preprocessing Module

The preprocessing module cleans and prepares the collected datasets by handling missing values, removing inconsistencies, and converting the data into structured time-series format. This module also performs data normalization and transformation operations required for forecasting analysis.

3. Forecasting Engine

The forecasting engine is the core analytical component of the system. It applies statistical forecasting models such as ARIMA and Facebook Prophet to analyze historical price patterns and generate predictions about future ingredient prices.

4. Volatility Detection Module

This module evaluates prediction confidence intervals and identifies potential supply chain volatility events. If predicted price ranges indicate a significant increase or instability, the system generates alerts that can be communicated to users.

5. Prediction Service API

The prediction results are delivered through a RESTful API implemented using FastAPI. This service allows external systems to retrieve predicted price values, volatility indicators, and historical trend data.

6. Visualization and Decision Support Interface

The final component of the architecture is the user interface layer, which displays forecasting results to business owners through charts, graphs, and alerts. This interface allows users to monitor ingredient price trends and make proactive purchasing decisions.

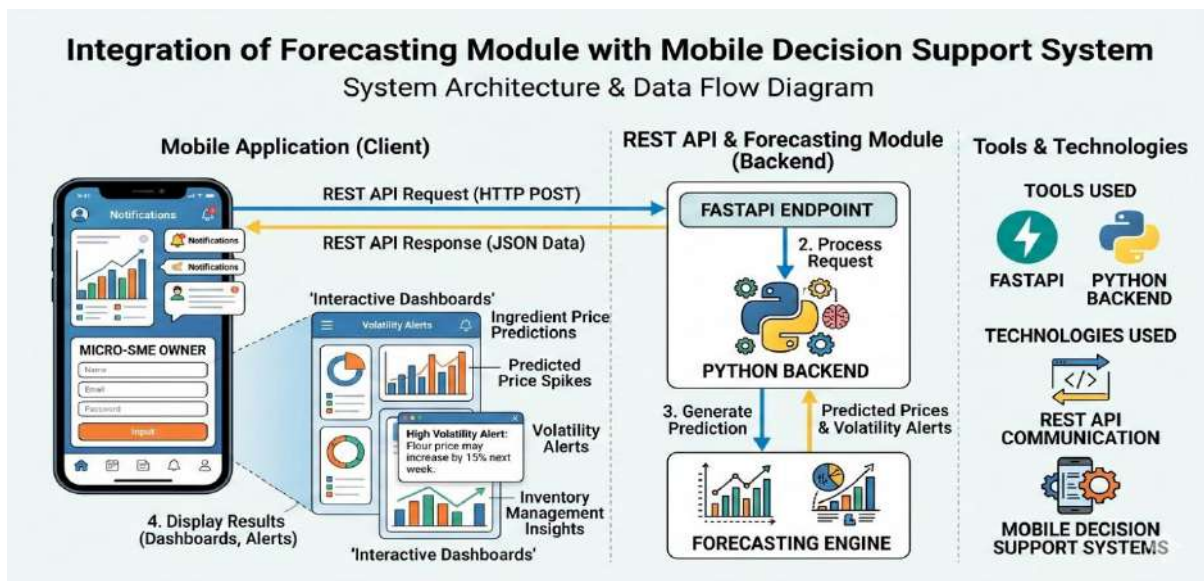


Figure 6

4.3 Component Interactions

The interaction between system components follows a structured workflow that ensures efficient data processing and forecasting analysis.

First, the data collection module retrieves historical commodity price datasets from external data sources. These datasets are passed to the data preprocessing module, where the data is cleaned, validated, and structured into a time-series format.

Once preprocessing is completed, the structured dataset is forwarded to the forecasting engine. The forecasting engine processes the data using ARIMA and Prophet models to generate predictions of future ingredient prices.

The predicted results are then evaluated by the volatility detection module, which analyzes prediction confidence intervals to identify potential price instability or supply chain volatility. If the system detects significant predicted price increases, an alert signal is generated.

Finally, the results are delivered to external applications through the prediction service API. The mobile or web-based decision-support interface retrieves this information and presents it to users in the form of visual charts, trend graphs, and predictive alerts.

4.4 Data Flow Explanation

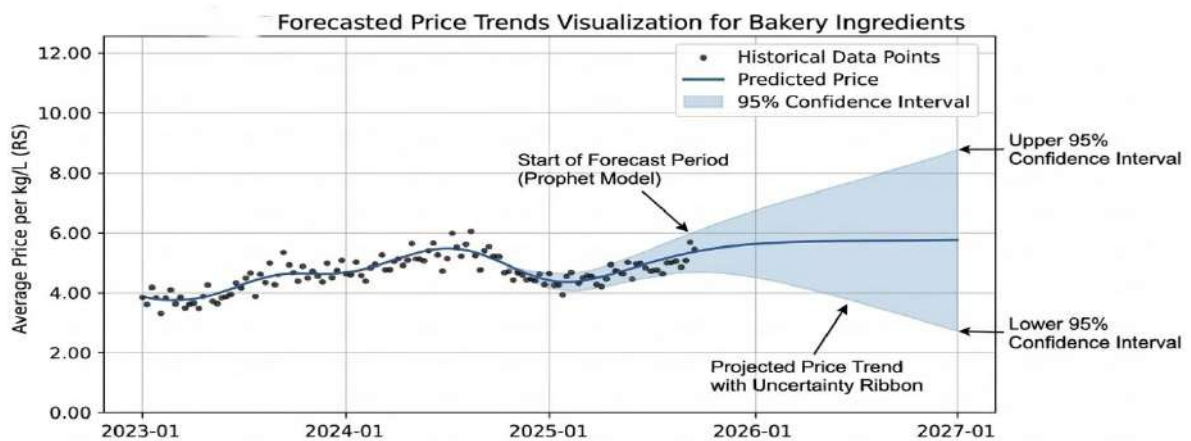
The data flow within the system begins with the acquisition of historical ingredient price datasets. These datasets typically contain records of commodity prices over time and serve as the primary input for the forecasting module.

The raw data is first processed by the preprocessing module, which performs operations such as data cleaning, time indexing, and normalization. After preprocessing, the data is structured as a time-series dataset containing chronological price observations.

The structured dataset is then fed into the forecasting engine, where statistical forecasting models analyze historical trends and generate predicted price values for future time periods. Along with predicted prices, the system also calculates confidence intervals that represent the uncertainty range of the predictions.

These forecasting results are then passed to the volatility detection module, which evaluates whether predicted price fluctuations exceed predefined thresholds. If potential price increases are detected, the system generates alerts indicating possible supply chain volatility.

The final prediction outputs are exposed through the API service and delivered to the decision-support interface, where users can view price forecasts and volatility alerts through graphical dashboards.



Forecasted Price Trends Visualization for Bakery Ingredients. This visualization compares historical average prices (black dots) with predicted values (blue line) and future confidence intervals (shaded area), created using the Prophet forecasting model with Python, Matplotlib, and Seaborn. Prices are in RS. Predictive insights show expected price movements and uncertainty.

Figure 7

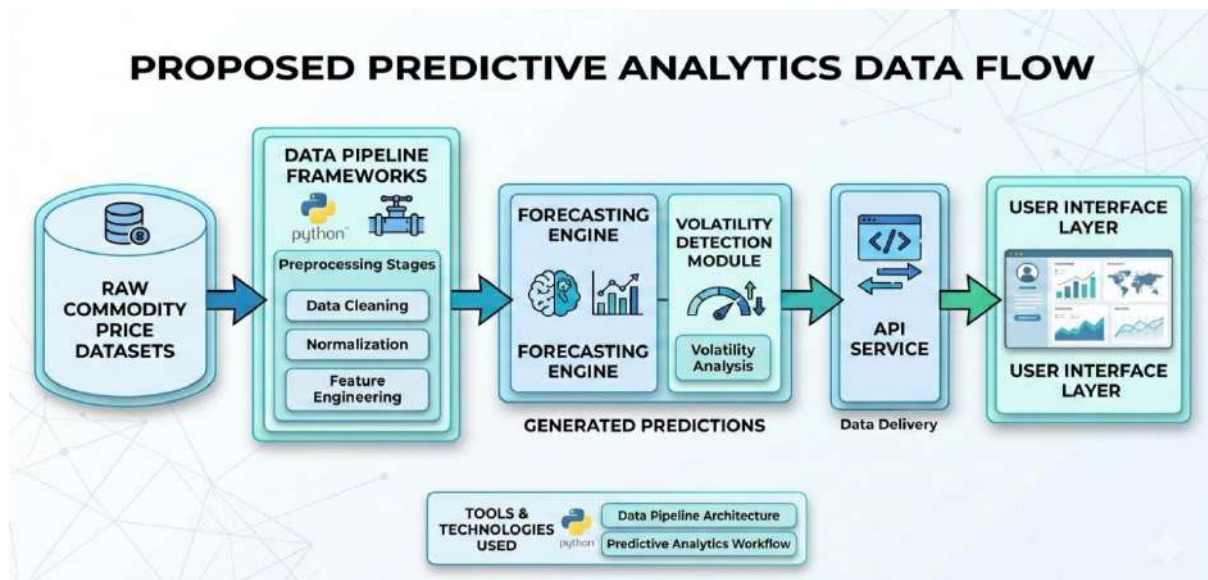


Figure 8

4.5 Scalability and Security Considerations

The proposed system architecture is designed with scalability and security considerations to ensure reliable operation in real-world environments.

From a scalability perspective, the forecasting module can be deployed as a cloud-based microservice that supports multiple prediction requests simultaneously. The use of FastAPI allows the system to handle high-performance API requests efficiently while maintaining low latency for prediction responses. Additionally, the modular architecture allows the forecasting models to be updated or extended with additional analytical techniques in the future.

The system also incorporates several security considerations. API endpoints will implement authentication mechanisms to prevent unauthorized access to forecasting services. Data communication between system components can be secured using encrypted communication protocols such as HTTPS.

Furthermore, since the system primarily uses publicly available commodity price datasets and does not process sensitive personal data, the risk of data privacy violations is minimal. Nevertheless, the system will follow responsible data management practices to ensure that all datasets are handled ethically and securely.

By incorporating scalability and security considerations into the architecture design, the proposed system can provide reliable predictive analytics services that support small business decision-making while maintaining system performance and data integrity.

5. USER REQUIREMENTS

5.1 Stakeholder Identification

Several stakeholders are involved in or affected by the proposed predictive forecasting system. Each stakeholder interacts with the system differently and has specific expectations regarding system functionality and performance.

1. Micro-SME Business Owners (Primary Users)

Micro-SME business owners, particularly bakery operators and small food production entrepreneurs, represent the primary users of the forecasting system. These users rely heavily on raw ingredients whose prices fluctuate frequently. The forecasting system will assist them in predicting future price trends, enabling them to make proactive purchasing decisions and maintain stable profit margins.

2. Inventory and Procurement Managers

In some small businesses, inventory or procurement personnel are responsible for managing ingredient purchases and stock levels. These stakeholders can use the forecasting module to monitor predicted price changes and plan procurement activities more effectively.

3. System Administrators and Developers

System administrators and developers are responsible for maintaining and updating the forecasting module. They ensure that the prediction models remain operational, manage system updates, and monitor system performance.

4. Academic Researchers and Analysts

Researchers and analysts may also interact with the system for evaluating forecasting models and analyzing supply chain volatility patterns. Their feedback may help improve the forecasting accuracy and performance of the system.

5.2 Functional Requirements

Functional requirements describe the specific capabilities that the predictive forecasting module must provide in order to support users in analyzing ingredient price trends and supply chain volatility.

FR1 – Data Input and Dataset Management

The system shall allow the ingestion of historical ingredient price datasets obtained from public data sources and commodity price repositories.

FR2 – Data Preprocessing

The system shall perform preprocessing operations on collected datasets, including data cleaning, normalization, missing value handling, and transformation into time-series format.

FR3 – Time-Series Forecasting Model Execution

The system shall implement forecasting algorithms such as ARIMA and Facebook Prophet to analyze historical price data and generate predictions for future ingredient prices.

FR4 – Prediction Output Generation

The system shall generate predicted price values for specified time periods, including confidence intervals that represent the uncertainty range of the forecasts.

FR5 – Supply Chain Volatility Detection

The system shall analyze predicted price trends and identify potential supply chain volatility events when predicted price changes exceed predefined thresholds.

FR6 – Alert and Notification Mechanism

The system shall generate alerts when predicted price increases or volatility patterns are detected, enabling users to take proactive purchasing actions.

FR7 – API-Based Prediction Access

The system shall provide a RESTful API through which external applications, such as mobile decision-support systems, can retrieve forecasting results and volatility alerts.

FR8 – Visualization Support

The system shall provide structured prediction data that can be visualized through dashboards, charts, and graphs in the user interface layer.

5.3 Non-Functional Requirements

Non-functional requirements describe the quality attributes and performance expectations of the predictive forecasting module.

Performance Requirements

The system should generate forecasting results within a reasonable response time to ensure timely access to predictive insights. Prediction requests processed through the API should return results with minimal latency.

Scalability Requirements

The architecture should support scalability to allow the forecasting module to handle multiple prediction requests simultaneously. The system should also be capable of incorporating additional forecasting models or larger datasets in the future.

Reliability Requirements

The forecasting system should maintain stable performance during continuous operation. The system should also include mechanisms to handle unexpected errors or missing data without interrupting prediction services.

Security Requirements

API endpoints should implement authentication and secure communication protocols to prevent unauthorized access. All data exchanges between system components should be protected using encrypted communication methods such as HTTPS.

Usability Requirements

The forecasting results should be presented in a format that is easy for business users to understand. Prediction outputs should be compatible with visualization tools that display ingredient price trends clearly.

Maintainability Requirements

The system architecture should follow modular design principles to ensure that forecasting models and system components can be updated or improved without affecting the entire system.

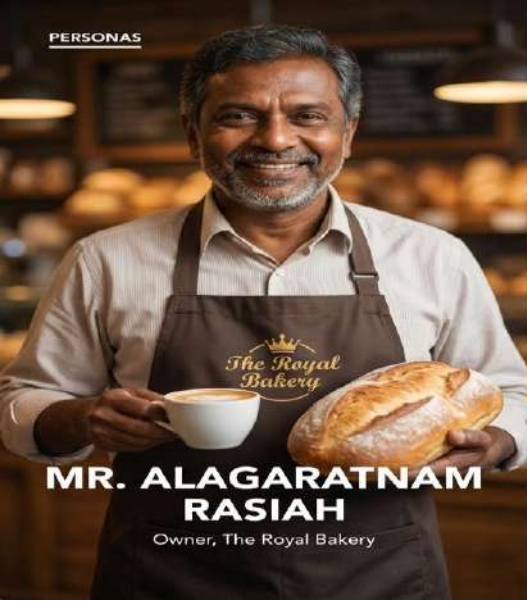
6. COMMERCIALIZATION PLAN

6.1 Target Market and Customer Persona

The primary target market for the proposed forecasting system includes micro and small-scale enterprises that depend heavily on raw material procurement, particularly businesses operating in the food production and retail sectors. And examples of potential target users include:

- Small bakeries

PERSONAS



MR. ALAGARATNAM RASIAH
Owner, The Royal Bakery

ABOUT

Mr. Rasiah owns a small neighborhood bakery. He manages raw material purchasing and production, creating bread, buns, and cakes daily. He seeks to stabilize operations and predict cost changes changes. (Text is summarized to fit style).

BUSINESS: The Royal Bakery
INDUSTRY: Small Bakery
KEY INGREDIENTS: Wheat Flour, Sugar, Butter, Yeast
MANAGEMENT ROLE: Purchasing & Production

GOALS

- Maintain stable profit margins despite cost
- Purchase raw materials at favorable prices
- Purchase raw materials at favorable prices
- Improve financial planning for operations

CHALLENGES

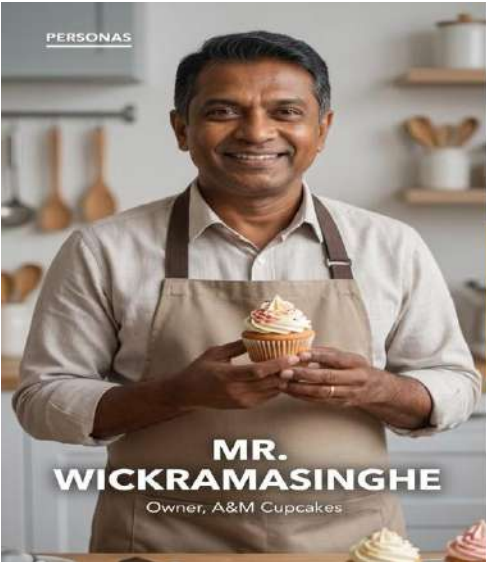
- Frequent changes in flour and sugar prices
- Limited access to advanced forecasting tools
- Difficulty predicting future ingredient price changes

SYSTEM BENEFITS

- Accurate future ingredient price predictions
- Early alerts about possible price increases
- Timely purchasing for best prices
- Avoid unexpected cost increases

- Home-based food businesses

PERSONAS



MR. WICKRAMASINGHE
Owner, A&M Cupcakes

ABOUT

Mr. Wickramasinghe runs A&M Cupcakes, a home-based business specializing in customized cupcakes and desserts for events. He seeks to stabilize operations and predict cost changes. (A short summary like the original is used).

BUSINESS: A&M Cupcakes
INDUSTRY: Home-Based Baking
SPECIALIZATION: Customized Cupcakes & Desserts
ROLE: Owner/Baker

GOALS

- Control ingredient costs while maintaining product quality
- Plan ingredient purchases efficiently for upcoming orders
- Increase profitability while running a small-scale operation

CHALLENGES

- Small profit margins due to fluctuating ingredient costs
- Limited knowledge of market price trends
- Difficulty planning ingredient purchases for future orders

SYSTEM BENEFITS

- Predict ingredient price trends and identify volatility
- Enable purchase planning in advance
- Maintain stable pricing for products

- Cafés and restaurants



ABOUT

Mr. Silva owns Butter Boutique, a café and restaurant serving baked goods, coffee, and light meals. He manages the costs of large quantities of dairy products, flour, sugar, and cooking oil. He seeks to stabilize operations and maintain competitive pricing.

BUSINESS: Butter Boutique

INDUSTRY: Café and Restaurant

KEY INGREDIENTS: Dairy, Flour, Sugar, Cooking Oil

ROLE: Owner/Manager

GOALS

- Monitor ingredient cost trends for long-term planning
- Maintain competitive menu pricing
- Reduce operational cost risks caused by supply chain fluctuations

CHALLENGES

- Sudden price increases in dairy products and cooking oil
- Difficulty forecasting future supply costs
- Need to balance product pricing with operational expenses

SYSTEM BENEFITS

- Predictive analytics provide forecasted price trends for key ingredients
- Allow adjustments to procurement strategies
- Support data-driven menu pricing

- Small grocery stores



ABOUT

Mrs. Murugan manages Cafe Shaze, a small grocery store selling food ingredients and bakery products to local customers. The store purchases ingredients in bulk and resells them. (Text is summarized to fit style).

BUSINESS: Cafe Shaze

INDUSTRY: Small Grocery and Food Store

SPECIALIZATION: Food Ingredients, Bakery Products

ROLE: Owner/Manager

GOALS

- Predict wholesale ingredient price changes
- Manage inventory efficiently
- Avoid purchasing products during peak price periods

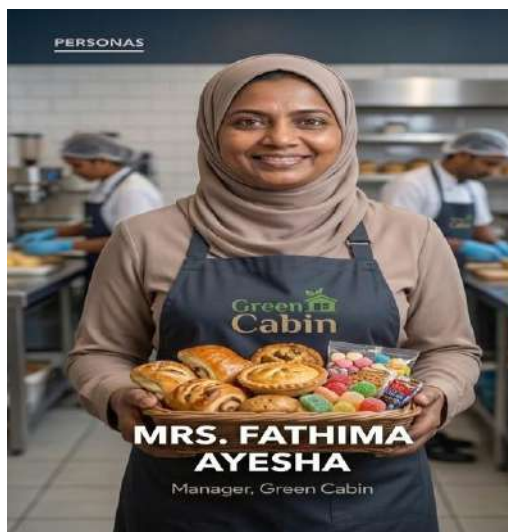
CHALLENGES

- Market price fluctuations affecting wholesale purchases
- Limited forecasting tools for small retailers
- Risk of overstocking or purchasing at high prices

SYSTEM BENEFITS

- Forecasting system enables predicted price trend analysis
- Identify supply chain volatility
- Help make informed purchasing decisions
- Manage inventory more effectively

- Food manufacturing micro-enterprises



ABOUT

Mrs. Ayesha is the manager of Green Cabin, a small food manufacturing business that produces baked snacks and confectionery products. The company depends on stable ingredient supply chains. (A summarized version, mirroring the style of

BUSINESS: Green Cabin

INDUSTRY: Small Food Manufacturing Enterprise

SPECIALIZATION: Baked Snacks & Confectionery

ROLE: Manager

GOALS

- Forecast ingredient price changes to improve production planning
- Maintain cost stability in manufacturing processes
- Reduce supply chain risks affecting raw materials

CHALLENGES

- Supply chain disruptions affecting ingredient availability
- Price volatility impacting production costs
- Difficulty planning procurement strategies

SYSTEM BENEFITS

- Predictive forecasting provides early warnings about potential price increases
- Company can secure raw materials in advance
- Maintain production efficiency

These businesses frequently purchase raw materials such as flour, sugar, cooking oil, dairy products, and other ingredients whose prices fluctuate due to supply chain disruptions and market conditions.

6.2 Value Proposition

The core value proposition of the proposed system is to empower micro-SMEs with predictive analytics capabilities that were previously available only to large enterprises.

The forecasting module provides several key benefits:

- Predicting future ingredient price trends
- Detecting supply chain volatility patterns
- Generating early alerts for potential price increases
- Supporting proactive procurement planning
- Reducing financial risks caused by fluctuating raw material costs

By providing predictive insights in an accessible and user-friendly format, the system enables small business owners to make informed decisions about when to purchase ingredients and how to manage inventory costs effectively.

6.3 Revenue Model

The forecasting system can be commercialized using a Software-as-a-Service (SaaS) model, where users access predictive analytics features through a subscription-based platform.

Several potential revenue models can be considered:

Subscription-Based Model

Small businesses pay a monthly or annual subscription fee to access forecasting services, predictive alerts, and analytics dashboards.

Freemium Model

A basic version of the system may be provided free of charge, offering limited forecasting features. Advanced forecasting analytics and alerts would be available through premium subscription plans.

API-Based Analytics Service

The forecasting engine may also be offered as an API service that other inventory management or business analytics platforms can integrate into their systems.

This flexible revenue structure allows the system to reach a broad range of users while maintaining a sustainable business model.

6.4 Cost Estimation

The development and operational costs of the forecasting system include several components.

Development Costs

Initial development requires time and effort from software developers and data scientists responsible for designing the forecasting models, backend services, and system integration.

Infrastructure Costs

If the system is deployed as a cloud service, infrastructure costs may include cloud computing resources, server hosting, and database storage.

Maintenance and Support Costs

Regular system updates, model retraining, and technical support services will also contribute to operational costs.

Since the forecasting system relies on open-source technologies such as Python, Pandas, Statsmodels, and Prophet, the software licensing cost is minimal, which helps reduce the overall operational cost of the system.

6.5 Pricing Strategy

The pricing strategy should be designed to ensure affordability for micro-SMEs while maintaining financial sustainability for the service provider.

A tiered subscription pricing model may be implemented:

- **Basic Plan** – Provides limited forecasting insights and historical trend analysis for small businesses.
- **Standard Plan** – Includes ingredient price predictions, volatility alerts, and visualization dashboards.
- **Premium Plan** – Offers advanced analytics, longer prediction horizons, and integration with external inventory systems.

Affordable pricing ensures that the solution remains accessible to micro-SMEs, which often operate with limited financial resources.

6.6 Competitive Advantage

- The proposed forecasting system offers several advantages compared to existing supply chain analytics platforms.
- First, the system is specifically designed for micro-SMEs, whereas most existing forecasting platforms target large enterprises.
- Second, the system is lightweight and cost-effective, making it accessible to small businesses that cannot afford enterprise-level ERP systems.
- Third, the system focuses specifically on predicting raw material price fluctuations, rather than only tracking inventory or demand forecasting.
- Finally, by using time-series forecasting models optimized for small datasets, the system can generate meaningful predictions even when historical data is limited.
- These features create a unique competitive advantage that positions the system as a practical predictive analytics solution for small business environments.

6.7 Intellectual Property Considerations

The intellectual property (IP) considerations of the proposed system focus on protecting the software design, forecasting algorithms integration, and system architecture developed during the research.

While the forecasting algorithms themselves (such as ARIMA and Prophet) are publicly available statistical methods, the integration of these models into a lightweight predictive analytics platform tailored for micro-SMEs represents a novel system implementation.

The following intellectual property protection strategies may be considered:

- Copyright protection for the developed software code
- Protection of system architecture and implementation design
- Trademark protection for the platform name and branding
- Licensing agreements if the forecasting API is provided to external platforms

By securing appropriate intellectual property rights, the system can be further developed into a commercial product while protecting the originality of the implementation.

7. BUDGET AND JUSTIFICATION

The development of the predictive supply chain volatility forecasting module requires a combination of hardware resources, software tools, cloud infrastructure, and human effort. This section outlines the estimated budget required to support the development and deployment of the forecasting system. Since the proposed solution relies heavily on open source technologies and publicly available datasets, the overall cost of implementation remains relatively low compared to traditional enterprise analytics systems.

The estimated budget focuses primarily on the computational resources required to develop and test the forecasting models, cloud infrastructure needed for system deployment, and the human effort required for system development and maintenance.

7.1 Hardware Costs

The forecasting module requires a standard computing environment capable of performing data preprocessing, statistical modeling, and time-series forecasting. The development and testing of the predictive analytics models will be conducted using a personal computer or workstation with moderate computational capacity.

A typical development workstation should include a multi-core processor, sufficient system memory, and adequate storage capacity for handling historical datasets and model training tasks. For this project, the recommended minimum hardware specifications include:

Table 5: Hardware Costs

Item	Quantity	Estimated Cost (LKR)	Purpose
Laptop / Development Computer	1	LKR 300,000	Used for data preprocessing, forecasting model development, and system testing
External Storage Device	1	LKR 20,000	Backup storage for historical ingredient price datasets and forecasting results
Internet Router / Connectivity	1	LKR 15,000	Required for dataset collection, cloud access, and API communication testing
Total Hardware Cost: LKR 335,000			

Since most forecasting computations rely on statistical models rather than deep learning algorithms, high-performance graphical processing units (GPUs) are not required. Therefore, the hardware cost remains relatively low and can typically be covered using existing academic laboratory equipment or personal development devices.

7.2 Software Licenses

One of the advantages of the proposed forecasting module is the use of open-source software technologies, which significantly reduces development costs.

The primary software tools used in this research include:

Table 6: Software Licenses Cost

Software Tool	License Type	Estimated Cost (LKR)	Purpose
Python Programming Language	Open Source	LKR 0	Main programming environment used for system development
Pandas & NumPy Libraries	Open Source	LKR 0	Data preprocessing and numerical analysis
Statsmodels Library	Open Source	LKR 0	Implementation of ARIMA forecasting model
Facebook Prophet Library	Open Source	LKR 0	Time-series forecasting and trend analysis
FastAPI Framework	Open Source	LKR 0	Backend API development
Matplotlib & Seaborn	Open Source	LKR 0	Visualization of forecasting results
Total Software Cost: LKR 0			

These technologies are freely available under open-source licenses and do not require paid subscriptions. As a result, there are no significant licensing costs associated with the development of the forecasting module.

However, optional tools such as professional development environments or advanced data analytics platforms could be used to improve development productivity. If required, tools such as JetBrains PyCharm Professional Edition or cloud-based development platforms could be considered.

7.3 Cloud Services

Cloud infrastructure may be required to deploy the forecasting module as a scalable backend service accessible through the internet. Cloud hosting platforms provide computational resources, storage, and networking capabilities required to operate the prediction service API.

Examples of potential cloud service providers include:

- Amazon Web Services (AWS)
- Google Cloud Platform (GCP)
- Microsoft Azure

The forecasting module can be deployed using lightweight cloud instances with moderate computing resources. Since the forecasting system processes relatively small datasets and performs periodic predictions rather than continuous high volume processing, the cloud resource requirements remain modest.

The estimated cloud costs may include:

- Virtual server hosting
- Data storage services
- API hosting infrastructure
- Network bandwidth usage

Table 7: Cloud Services Cost

Cloud Service	Estimated Cost (LKR)	Purpose
Virtual Server Hosting	LKR 15,000	Hosting the forecasting module backend API
Cloud Storage	LKR 8,000	Storage of forecasting datasets and prediction results
API Hosting Infrastructure	LKR 10,000	Deployment of the prediction service API
Network Bandwidth	LKR 7,000	Data transfer and API communication
Estimated Cloud Cost: LKR 40,000		

For academic or prototype deployments, many cloud providers offer free-tier services that can support the initial development and testing phases. This helps minimize infrastructure costs during the research stage.

7.4 Human Effort Cost

Human effort represents one of the most significant components of the overall project cost. The development of the forecasting module requires time and expertise in data science, software engineering, and system integration.

The key development tasks include:

Table 8: Human Effort Cost

Role	Estimated Hours	Hourly Rate (Estimated)	Total Cost
Frontend & Computer Vision Developer (Member 1)	200 hours	LKR 3,200	LKR 640,000
NLP & Data Processing Developer (Member 2)	180 hours	LKR 3,200	LKR 576,000
Predictive Analytics Developer (Member 3)	180 hours	LKR 3,200	LKR 576,000
Backend & Financial Logic Developer (Member 4)	200 hours	LKR 3,200	LKR 640,000
Total Human Effort Cost: LKR 2,432,000			

In an academic research context, the development work is typically performed by students as part of their research project. Therefore, the direct financial cost of human effort may be minimal during the research phase.

However, if the system is later commercialized, additional resources may be required for software development, system maintenance, customer support, and model improvements.

Overall, the proposed forecasting system is designed to be cost-effective and accessible, allowing micro-SMEs to benefit from predictive analytics without requiring large financial investments. The use of open-source technologies and scalable cloud infrastructure ensures that the solution remains economically feasible while delivering valuable predictive insights.

Total Estimated Project Budget:

Table 9: Total Estimated Project Budget

Category	Estimated Cost
Hardware	LKR 335,000
Software	LKR 0
Cloud Services	LKR 40,000
Human Effort	LKR 2,432,000
Total Estimated Project Budget: LKR 2,807,000	

8. WORK BREAKDOWN STRUCTURE (WBS)

Table 10: Work Breakdown Structure for Predictive Analytics Component

Task ID	Task Description	Tools / Technology	Duration
T1	Research on Time-Series Forecasting	Research Papers, Literature Review	2 Weeks
T2	Dataset Collection	Python, Web Data Sources	1 Week
T3	Data Cleaning & Preprocessing	Pandas, NumPy	2 Weeks
T4	ARIMA Model Implementation	Python, Statsmodels	2 Weeks
T5	Prophet Model Implementation	Python, Prophet Library	2 Weeks
T6	Model Evaluation	Scikit-learn Metrics	1 Week
T7	API Development	FastAPI	2 Weeks
T8	Integration with Mobile System	REST API	2 Weeks
T9	Testing and Validation	Python Testing Tools	1 Week

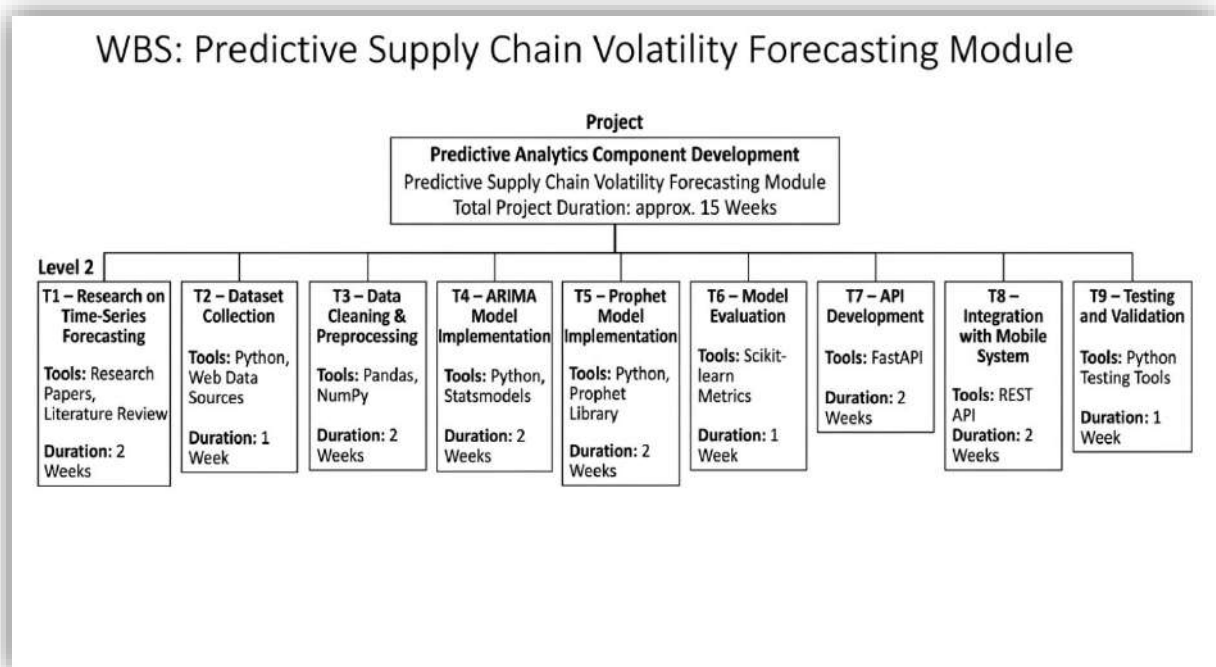


Figure 9

9. GANTT CHART

Smart-Cost: Hybrid ARIMA-Prophet-LSTM Volatility Forecasting with Agentic News Fusion for Commodity Price Prediction in Resource-Constrained SMEs

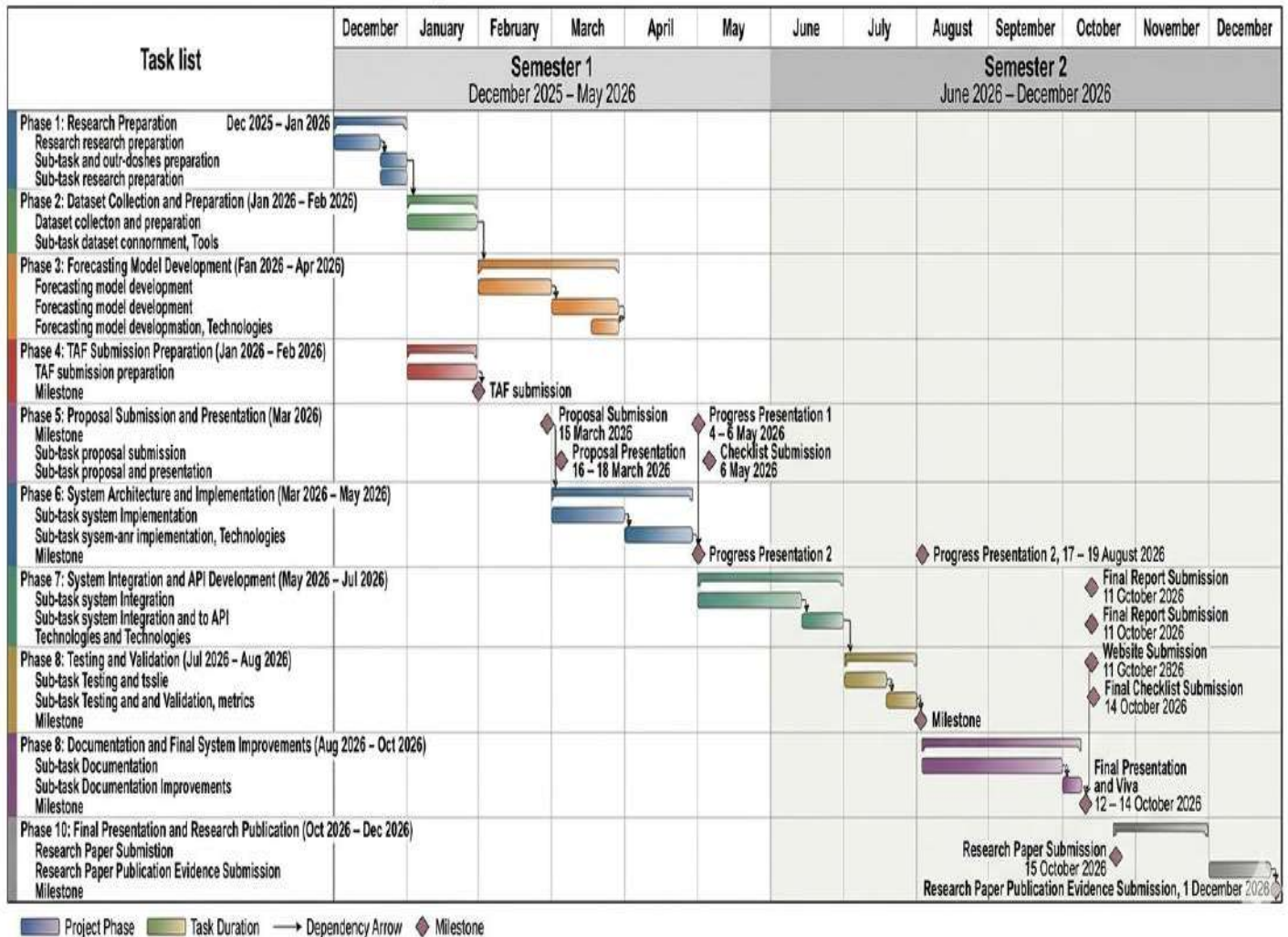


Figure 10

10. REFERENCES LIST

- [1] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, 2nd ed. San Francisco, CA, USA: Holden-Day, 1976.
- [2] G. Zhang, B. E. Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks: The state of the art,” *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.
- [3] S. J. Taylor and B. Letham, “Forecasting at scale,” *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.
- [4] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [5] SAP SE, “SAP Integrated Business Planning for Supply Chain,” SAP Official Documentation. [Online]. Available: <https://www.sap.com/products/scm/integrated-business-planning.html>
- [6] Zoho Corporation, “Zoho Inventory: Inventory management software for growing businesses.” [Online]. Available: <https://www.zoho.com/inventory/>
- [7] T. M. Choi, S. W. Wallace, and Y. Wang, “Big data analytics in operations management,” *Production and Operations Management*, vol. 27, no. 10, pp. 1868–1883, 2018.
- [8] G. Wang and H. Zhang, “Data-driven forecasting and decision making for small and medium enterprises,” *Journal of Business Research*, vol. 98, pp. 123–135, 2019.
- [9] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, Australia: OTexts, 2021.
- [10] R. J. Hyndman and Y. Khandakar, “Automatic time series forecasting: The forecast package for R,” *Journal of Statistical Software*, vol. 27, no. 3, pp. 1–22, 2008.

11.APPENDICES

This section presents supplementary materials that support the design, development, and validation of the predictive supply chain volatility forecasting module proposed in this research. The appendices include survey instruments used to understand the needs of micro-SME business owners, system diagrams illustrating the architecture and workflow of the forecasting system, preliminary user interface mockups of the mobile decision support platform, a risk analysis identifying potential technical and operational risks, and ethical clearance documentation related to the use of publicly available datasets. These materials provide additional context for the proposed research methodology and demonstrate how the forecasting module can be implemented and evaluated in real-world micro-SME environments.

A. Survey Instruments

To better understand the challenges faced by micro and small-scale enterprises (micro-SMEs) in managing ingredient price volatility, a preliminary survey instrument was developed. The purpose of the survey is to collect insights from business owners operating in sectors such as bakeries, cafés, home-based food businesses, grocery stores, and small food manufacturing enterprises. These businesses frequently rely on raw materials such as flour, sugar, butter, cooking oil, and dairy products, whose prices are highly volatile due to supply chain disruptions and market fluctuations.

The survey aims to identify current practices used by micro-SME owners to monitor ingredient price changes and determine whether predictive analytics tools could assist them in making better procurement decisions. The results of this survey help guide the design of the forecasting module and ensure that the proposed system addresses real-world operational challenges.

Sample Survey Questions

1. What type of food-related business do you operate?
 - Small bakery
 - Home-based food business
 - Café or restaurant
 - Grocery store
 - Food manufacturing micro-enterprise
2. Which ingredients do you purchase most frequently?
 - Wheat flour
 - Sugar
 - Butter or dairy products
 - Cooking oil
 - Other ingredients
3. How often do you experience sudden price changes in these ingredients?
 - Very frequently
 - Occasionally
 - Rarely
4. How do you currently monitor ingredient price changes?
 - Manual observation

- Supplier updates
- Market monitoring
- Inventory management software

5. Would a predictive system that forecasts ingredient price changes help your business plan purchases more effectively?

- Yes
- No
- Not sure

6. How far in advance would price predictions be useful for your business?

- 1 week
- 1 month
- 3 months

B. System Diagrams

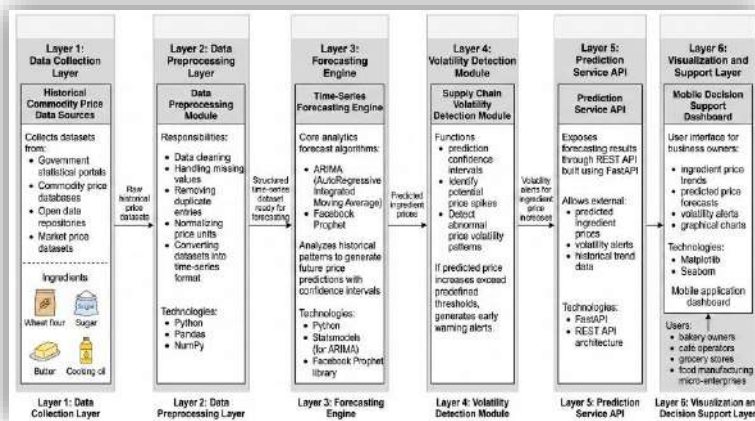


Figure 11

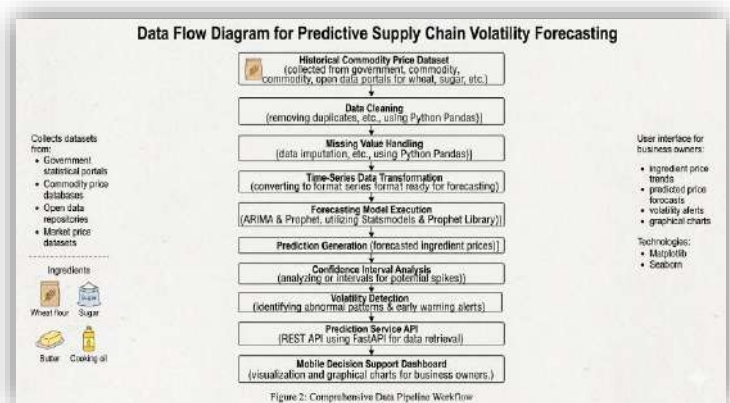


Figure 12

C. UI Mockups

This appendix presents preliminary user interface mockups for the mobile decision-support application that integrates with the forecasting module. The purpose of the mobile interface is to present predictive insights in a simple and understandable format for micro-SME business owners who may not have technical expertise in data analytics.

Dashboard Interface

The dashboard provides an overview of ingredient price trends and predictive insights. The interface displays:

- Current ingredient price trends

- Forecasted price predictions
- Supply chain volatility alerts
- Graphical visualization of price changes over time

The dashboard allows users to quickly monitor market trends and identify potential price increases.

Ingredient Forecast Screen

This screen provides detailed forecasting information for individual ingredients. Users can select a specific ingredient such as wheat flour, sugar, or cooking oil and view the following information:

- Historical price trends
- Predicted price curve
- Forecast confidence intervals
- Forecast time horizon (weekly or monthly)

Notification and Alert Interface

This interface provides early warning alerts when the forecasting model predicts significant price increases.

Example alert message:

“Predicted increase in wheat flour price within the next two weeks. Consider purchasing ingredients early to reduce cost impact.”

These UI mockups demonstrate how predictive analytics results can be presented to business users in a clear and accessible format.

D. Risk Analysis

This appendix identifies potential risks associated with the development and deployment of the predictive forecasting module and proposes mitigation strategies to reduce their impact.

Risk	Description	Mitigation Strategy
Data Quality Issues	Historical datasets may contain missing or inconsistent values	Apply preprocessing techniques such as interpolation and data cleaning
Forecasting Inaccuracy	Prediction models may produce inaccurate results if market patterns change significantly	Retrain forecasting models regularly using updated datasets
Supply Chain Disruptions	External economic events may cause sudden market changes	Incorporate confidence intervals and update predictions periodically
System Downtime	Cloud infrastructure or API services may experience interruptions	Deploy system on reliable cloud platforms with backup services
Limited Dataset Availability	Small businesses may have limited historical data	Combine public datasets with local market data

11.5 Ethical Clearance Documents

SLIIT FACULTY OF COMPUTING
Research Ethics Review Form

All students undertaking research in the Faculty of Computing must fill in and submit this form to the respective coordinator / supervisor as instructed by the program head.

SECTION 1

1. Project/research type

☒ B.Sc. ☐ M.Sc. ☐ MPhil/PhD

2. Project/research details

Project ID (if any): R25-IT-026
Research Title: Smart Cost : Smart Supply Chain Cost Analytics and Profit Prediction System for Micro-SMEs

3. Student details (if the project is a group project, insert the details of group leader first)

Student ID	Name	Email	Phone
IT22544872	Kodithuwakku K.A.C.H	IT22544872@my.sliit.lk	0760152847
IT22151810	Perera N.K.M	IT22151810@my.sliit.lk	0769864100
IT22162250	Jayasingha D.D.M	IT22162250@my.sliit.lk	0773055133
IT22152664	Dasanayaka M.P.J.R	IT22152664@my.sliit.lk	071200554

4. Supervisor details

Supervisor Name: Mr. Deemantika Siriwardana
Co-supervisor Name (if any): Ms. Aysha Wijesooriya
External supervisor (if any):

SECTION 2: DATA COLLECTION AND STORAGE

Read the following table before answering the questions.

Data Source	Description	Data Collection Methods
Primary Data	Primary data collection happens when researchers obtain information directly from the original sources.	Survey/Interviews/ Observation/ Focus groups/ New Images etc.
Secondary Data	Datasets already gathered from previous research. This data is usually already analyzed and put into context. Sources of secondary data can include:	Books / Past research papers / Annual reports / Newspapers / Reports / Websites / Online Databases etc.

Page 01

Provide details of all data required for your project. (Indicate what data will be required for each component, where the data is available, and how you plan to collect that data.)

Research Component	Data required. (e.g. Images of leaves)	Amount of Data Required	Data Source (Primary /Secondary)	Collection Method
Adaptive Thermal Thermal Restoration for OCR Based Inventory Digitization	Thermal receipt images from Supermarkets, and small food businesses	2000-3000 receipt images (Open, Valid, damaged)	Primary Data	Direct collection from shops, mobile camera, capture
Semantic Inventory Normalization Engine	Ingredient names, abbreviations, quantities, and measurement units from receipts	2000-3000 receipt images (Open, Valid, damaged)	Primary + Secondary Data	Collected from collected receipts, manual dataset creation
Hybrid Predictive Supply Chain Volatility Forecasting	Historical ingredient price data (Rice, sugar, oil, onion etc.)	2-3 years historical price records	Secondary Data	Government statistics websites, Kaggle datasets, other datasets
Real-time Receipts & Dynamic Forecasting System	Serialized sales transactions and ingredient purchase records	200-300 simulated business transactions	Primary Data	Generated test dataset and manual data entry

5. DATA COLLECTION: If you require primary data, do you intend to collect data from following groups?

☐ Children or young people
☐ Pregnant women
☐ Individuals with cognitive disabilities
☐ People suffering from specific diseases
☐ Animals
☒ Certain group of people involved in an organization

6. DATA COLLECTION: If you are collecting information related to people, in what form will be the data collected?

☒ Individually identifiable organization
☒ Re-identifiable* Data indirectly leads to the identification of a person or organization (even through direct identifiers are removed)
☐ Non-Identifiable Data does not lead to the identification of a specific individual
*Anonymized data from which the individual could be discovered

7. DATA COLLECTION: Does your project possibly cause/involve any of the following?

☐ Physical stress/distress, or discomfort to participants
☐ Psychological/mental stress/distress, or discomfort to participants
☒ Collection of sensitive/confidential data of individuals
☒ Collection of confidential data of organizations
☐ Development of a product harmful for a particular group or society

Page 02

8. DATA COLLECTION: Explain the procedures you will be following to ensure privacy of participants or organizations during data collection, data analysis and publishing of results.

☒ Before the data collection, participants will be given an information sheet to create awareness
☒ Data will be collected only from the participants who have given consent through a consent form
☒ Personally identifiable data will not be collected
☒ If identifiable data is present in the data set, data will be either anonymized or coded

Additional Measures:
The research will primarily collect receipt images and publicly available price datasets, which do not contain personal information. If any business-related information appears on receipts (such as shop names or addresses), it will be removed or anonymized during preprocessing. All collected data will be used strictly for academic research purposes.

9. DATA STORAGE: If you are collecting information related to people, in what form will the data be stored?

☐ Individually identifiable ☐ Re-identifiable ☒ Non-Identifiable

10. DATA STORAGE: Where will data be stored and what measures will be taken to ensure security of data?

☐ Paper-based data stored in locked filing cabinets within SLIIT.
☒ Data stored in hard-drives and computers will be protected by password and encryption.
☒ Data will be anonymized and will not be re-identifiable.
☐ After the completion of the project, all data will be permanently deleted from all student devices. Only a single copy will be securely stored within a repository in the institution.

Additional measures:
All research data including receipt images, datasets, and experimental results will be securely stored in password-protected computers and cloud storage. Access to the data will be limited only to the research team members and supervisors. The datasets will be anonymized and used strictly for academic research purposes.

11. How would you rate the overall risk associated with the research? (These risks may occur when the collected data is exposed to unauthorized parties, etc.)

☐ High Risk ☒ Low Risk ☐ Negligible Risk

Page 03

12. Declaration of the participants.

I/we certify that the details given above are accurate and agree to adhere to the conditions of approval specified by the panel. Furthermore, I/we will inform the FoC Research Project coordinator of any deviation from the specified details.

Student ID	Name	Signature
IT22544872	Kodithuwakku K.A.C.H	
IT22151810	Perera N.K.M	
IT22162250	Jayasingha D.D.M	
IT22152664	Dasanayaka M.P.J.R	

13. Approval of the Supervisor

Supervisor Recommendation (please select one option): ☐ Ethics committee Approval NOT Required ☐ Ethics committee Approval Required

If approval required, clearly specify the reasons:

Supervisor Name: Supervisor Signature: Date:

**** Please note that if the supervisor has recommended that ethics committee approval required, student should fill the 'Ethics Review Form' of the SLIIT ERC and submit it to the faculty Ethics Sub-committee.**

Page 04