

A Smart IoT-Driven Platform for Sustainable Dry Fish Processing in Sri Lanka

RP-026-079

Project Proposal Report

VIVIPEM L B R V – IT22639844

B.Sc. (Hons) Degree in Information Technology

Information Technology

Department of Information Technology

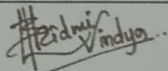
Sri Lanka Institute of Information Technology, Sri Lanka

March 2026

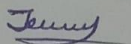
Declaration

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
VIVIPEM L B R V	IT22639844	

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

Name	Date	Signature
Mrs. Jenny Krishara	3/14/2025	

Abstract

Traditional dry fish production in Sri Lanka is a centuries-old livelihood that supports approximately 50,000 coastal processors and contributes significantly to the country's food security and export economy. However, the process remains almost entirely manual and weather-dependent, resulting in an estimated 20–30% post-harvest loss due to over-drying, under-drying, and spoilage from unexpected humidity changes. Processors currently rely on visual inspection and personal experience to judge when fish are adequately dried or at risk of spoilage, a subjective method that is inconsistent, non-scalable, and fails to catch early warning signs that sensors can detect. This research proposes a Machine Learning Based Predictive Analytics Module as part of a broader Smart IoT-Driven Platform for sustainable dry fish processing. The module ingests time-series data from IoT sensors (ESP32 with DHT22 temperature/humidity sensors and HX711 load cells for weight measurement) deployed, and performs two core predictions: (1) drying completion time estimation using regression models (Random Forest Regressor, Gradient Boosting Regressor), and (2) spoilage risk classification (low/medium/high) using Random Forest and Gradient Boosting classifiers. The module engineers domain-specific features from raw sensor streams including rolling averages, humidity trend slopes, weight loss rate, temperature variation, and sudden humidity jump events to capture the temporal dynamics that drive drying outcomes. The system is evaluated by comparing baseline models (Linear Regression, Logistic Regression) against ensemble methods using MAE, RMSE for regression and Accuracy, Precision, Recall, F1-score for classification. The prediction service is deployed via FastAPI and integrated with the platform's Flutter mobile dashboard, enabling processors to receive real-time drying progress estimates and early spoilage alerts directly on their phones.

Keywords: *Machine Learning, Internet of Things (IoT), Predictive Analytics, Dry Fish Processing, Spoilage Risk Prediction.*

Acknowledgement

I would like to express my sincere gratitude to my project supervisor, Ms. Jenny Krishara, for her continuous guidance, constructive feedback, and invaluable support throughout the development of this research. Her insightful comments and expertise in software engineering greatly contributed to improving the quality and direction of this project.

I am also thankful to my co-supervisor, Dr. Dinuka Wijendra, for providing valuable academic perspective and methodological guidance. Special appreciation goes to our external supervisor, Mr. Dinuth Rashmika Jayasekara, a dry fish business owner whose practical industry knowledge and firsthand understanding of the challenges in dry fish processing were essential in shaping the real-world relevance of this research.

My appreciation extends to my project team members, for their collaborative spirit, shared technical discussions, and mutual motivation during the research and development phases. I am also grateful to the academic staff of the Department of Information Technology at SLIIT for providing the necessary knowledge and resources.

Finally, I am deeply grateful to my family for their unwavering patience, support, and encouragement throughout my academic journey.

Table of Contents

Declaration	2
Abstract	3
Acknowledgement	4
Table of Contents	5
List of Tables	7
List of Figures	7
List of Abbreviations	8
List of Appendices	8
1. Introduction	9
1.1 Background and Context	9
1.2 Literature Review	11
1.2.1 IoT-Based Fish and Food Drying Systems	11
1.2.2 Machine Learning for Drying Time Prediction	11
1.2.3 Spoilage Prediction Using Sensor Data	12
1.2.4 Time-Series Analysis in Food Processing	12
Existing Products and Systems	13
1.3 Research Gap	15
2. Objectives	17
2.1 Main Objective	17
2.2 Specific Objectives	17
3. Methodology	18
3.1 Research Approach	18

3.2 Technical Methods.....	18
Data Collection	18
Feature Engineering.....	19
Model Training and Evaluation	20
3.3 Tools and Platforms	22
3.4 Data Collection and Ethical Considerations	22
3.5 Validation Strategy	23
4. High-Level System Architecture	24
4.1 System Overview	24
4.2 ML Prediction Pipeline	27
4.3 Data Flow and Integration	28
4.4 Scalability and Security	28
5. User Requirements.....	29
5.1 Stakeholder Identification.....	29
5.2 Functional Requirements	30
5.3 Non-Functional Requirements	31
6. Commercialization Plan.....	31
6.1 Target Market and Customer Persona.....	31
6.2 Value Proposition.....	32
6.3 Revenue Model	33
6.4 Pricing Strategy.....	34
6.5 Competitive Advantage	34
6.6 Intellectual Property Considerations.....	34
7. Budget and Justification.....	35
8. Work Breakdown Structure	36

9. Gantt Chart.....	37
References.....	38
Appendices.....	39
Appendix A – Survey Instruments (Problem Validation).....	39
Appendix B – Risk Analysis.....	40

List of Tables

<i>Table 1: Comparison of Existing Research Methods and the Proposed System</i>	<i>14</i>
<i>Table 2: Sensor Data Features</i>	<i>20</i>
<i>Table 3: Machine Learning Model Training Setup</i>	<i>21</i>
<i>Table 4: Tools and Platforms</i>	<i>22</i>
<i>Table 5: Model Validation Metrics and Evaluation Criteria</i>	<i>23</i>
<i>Table 6: Stakeholder Identification</i>	<i>29</i>
<i>Table 7: Functional Requirements.....</i>	<i>30</i>
<i>Table 8: Non-functional Requirements</i>	<i>31</i>
<i>Table 9: Budget Estimation</i>	<i>35</i>
<i>Table 10: Work Breakdown Structure</i>	<i>37</i>
<i>Table 11: Risk Analysis.....</i>	<i>40</i>

List of Figures

<i>Figure 1: Feature Engineering Pipeline.....</i>	<i>19</i>
<i>Figure 2: Feature Engineering Pipeline.....</i>	<i>19</i>
<i>Figure 3: High Level System Architecture Diagram</i>	<i>24</i>
<i>Figure 4: Individual System Diagram</i>	<i>25</i>
<i>Figure 5: ML Prediction Pipeline (6 Steps)</i>	<i>27</i>
<i>Figure 6: Project Gantt Chart</i>	<i>37</i>

List of Abbreviations

Abbreviation	Full Form
ML	Machine Learning
IoT	Internet of Things
API	Application Programming Interface
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
RF	Random Forest
GB	Gradient Boosting
DHT22	Digital Humidity and Temperature Sensor
HX711	Load Cell Amplifier Module
ESP32	Espressif Systems 32-bit Microcontroller
DSR	Design Science Research
FAO	Food and Agriculture Organization
NARA	National Aquatic Resources Research and Development Agency
LKR	Sri Lankan Rupees
SDG	Sustainable Development Goal
OSS	Open-Source Software
GDP	Gross Domestic Product

List of Appendices

Appendix A – Survey Instruments (Problem Validation)

Appendix B – UI Mockups Description

1. Introduction

1.1 Background and Context

Sri Lanka's fisheries sector plays an important role in the national economy, contributing approximately 1.3% to the Gross Domestic Product (GDP) and providing livelihoods for over 250,000 coastal fishing families [1]. Among fisheries products, dry fish (locally known as "karawala") is one of the most widely consumed value-added products and is also exported to markets in South Asia and the Middle East. According to estimates by the Food and Agriculture Organization (FAO), Sri Lanka produces approximately 40,000–50,000 metric tonnes of dried fish annually, with most production carried out by small-scale coastal processors using traditional sun-drying methods [1][2].

Despite its economic importance, the dry fish processing industry faces significant quality control challenges. Traditional drying practices depend entirely on environmental conditions such as temperature, humidity, and wind, which are unpredictable, particularly during monsoon transitions. Processors typically rely on experience, visual inspection, and manual touch to determine whether fish batches are adequately dried. This subjective approach frequently results in inconsistent product quality and post-harvest losses. Studies estimate that approximately 20–30% of processed fish is lost due to over-drying, under-drying, or spoilage caused by sudden humidity changes [2][3].

This issue affects multiple stakeholders across the dry fish value chain. Small-scale processors, estimated at around 50,000 families, experience direct financial losses when fish batches spoil or are rejected by buyers. Traders and wholesalers face uncertainty in product quality, leading to price disputes and inefficient supply chains. Consumers are exposed to potential food safety risks when improperly dried fish with high moisture content reaches the market, increasing the likelihood of bacterial growth, fungal contamination, and aflatoxin formation. Additionally, inconsistent quality limits Sri Lanka's export potential, as international markets increasingly require reliable quality standards and traceability systems [1].

Existing approaches for monitoring fish drying have several limitations. Most processors still rely on fully manual drying methods with no technological support. Some research initiatives

have introduced IoT-based monitoring systems that collect environmental data such as temperature and humidity; however, these systems mainly provide real-time sensor readings without predictive insights [4][5]. Industrial drying machines with automated controls are available but are too expensive for small-scale processors, typically costing between USD 10,000 and USD 50,000. Furthermore, while machine learning techniques have been applied in other agricultural drying domains, such as grain and fruit drying, their application to fish drying using IoT sensor data in tropical developing contexts remains limited [3][6][7].

To address these challenges, this research proposes a Machine Learning–Based Predictive Analytics Module as part of a Smart IoT-Driven Digital Platform for sustainable dry fish processing. The system uses IoT sensors to collect environmental data, including temperature, humidity, and fish weight, during the drying process. Using this data, machine learning models perform two key predictive tasks. First, a regression model estimates the remaining drying time required for a batch to reach safe moisture levels. Second, a classification model predicts spoilage risk levels (low, medium, high) based on environmental conditions and drying progress. These predictions allow processors to take preventive actions and make informed decisions during processing. Ensemble machine learning techniques such as Random Forest and Gradient Boosting are used to analyze engineered features derived from time-series sensor data. The prediction service is deployed through FastAPI and integrated into a Flutter-based mobile dashboard, enabling processors to receive real-time insights and alerts.

The proposed platform contributes to several United Nations Sustainable Development Goals (SDGs). By reducing post-harvest losses in fish processing, the system supports SDG 2: Zero Hunger. Improving income stability for small-scale processors aligns with SDG 8: Decent Work and Economic Growth. Minimizing waste through better quality control contributes to SDG 12: Responsible Consumption and Production, while improving processing efficiency can indirectly support SDG 14: Life Below Water by reducing pressure on marine resources.

From a research perspective, this work aligns with multiple research clusters, including Machine Learning and Data Analytics, through the application of supervised learning models to IoT time-series data; Internet of Things (IoT), through ESP32-based environmental monitoring systems; and Software Engineering, through the development of a modular, API-based prediction service integrated with mobile and backend systems..

1.2 Literature Review

This section reviews key scholarly and technical sources that form the theoretical foundation of the proposed predictive analytics module, organized by research domain.

1.2.1 IoT-Based Fish and Food Drying Systems

A study investigating post-harvest losses in Sri Lanka's traditional dry fish industry identified that inadequate drying control, particularly during monsoon season transitions, is a primary cause of quality degradation. The research reported that up to 25% of production may be lost to spoilage in certain coastal regions, highlighting the need for improved monitoring systems for the drying process [3].

A research demonstrated an IoT-based monitoring system for solar drying of agricultural products, using temperature and humidity sensors to track environmental conditions in real time. While the system successfully recorded drying parameters, it lacked predictive capability and only displayed current readings without estimating drying completion time or spoilage risk levels [4].

Further research proposed a smart fish drying system using Arduino-based sensors with a threshold-based alert mechanism. The system confirmed the feasibility of deploying low-cost IoT sensors in tropical drying environments; however, the approach relied on simple rule-based alerts rather than data-driven predictive models. These studies provide the technological foundation for integrating machine learning into IoT-based drying systems [5].

1.2.2 Machine Learning for Drying Time Prediction

Research applying machine learning models such as Artificial Neural Networks (ANN) and Random Forest Regression to agricultural drying processes demonstrated that ensemble models can effectively predict drying kinetics. The results showed that Random Forest models outperform traditional thin-layer drying models, particularly when temporal features such as rolling averages and trend slopes are used instead of instantaneous measurements. These findings inform the feature engineering strategy used in the proposed predictive module [6].

A study used Gradient Boosting algorithms to model fruit drying under varying environmental conditions. The results indicated that ensemble learning approaches significantly outperformed

traditional regression techniques, achieving 15–20% lower RMSE compared with Linear Regression and Support Vector Regression models. This evidence supports the use of ensemble methods such as Random Forest and Gradient Boosting in the proposed predictive framework [7].

1.2.3 Spoilage Prediction Using Sensor Data

A comprehensive review of predictive models for seafood spoilage detection identified that environmental variables such as temperature–humidity interactions and exposure duration above critical thresholds are strong indicators of microbial growth and spoilage risk. The review also emphasized that classification-based prediction models, which categorize risk levels, are more practical for operational decision-making compared to continuous spoilage indices. This finding supports the proposed system’s three-level spoilage risk classification (low, medium, high) [8].

Another study explored the relationship between time–temperature exposure and food quality degradation, introducing the Time-Temperature Indicator (TTI) framework. This framework demonstrates how cumulative exposure to unfavorable environmental conditions accelerates quality loss in food products. These principles inform the proposed system’s feature engineering approach, particularly the calculation of cumulative time above risk thresholds as a predictive indicator of spoilage [9].

1.2.4 Time-Series Analysis in Food Processing

Research applying machine learning to food safety monitoring using IoT time-series data showed that temporal feature engineering significantly improves prediction performance. By capturing patterns such as trends, periodicity, and rates of environmental change, prediction accuracy increased compared to models trained on raw sensor values. Random Forest models in this study achieved F1-scores above 0.85 for food safety risk classification, establishing a performance benchmark for the proposed system [10].

Another study focusing on fish drying kinetics demonstrated that the rate of weight loss is a more reliable indicator of drying progress than absolute weight measurements. This insight supports the use of derivative-based features such as weight loss rate, which will be calculated from load cell sensor data in the proposed predictive module [11].

Existing Products and Systems

- **Manual Processing (Status Quo):** The vast majority of Sri Lanka's dry fish processors use no monitoring technology. Drying completion is judged subjectively by appearance and texture, leading to significant variability and waste.
- **Basic Sensor Dashboards:** Research prototypes display real-time sensor readings on web or mobile dashboards but provide no predictive intelligence. They tell a processor what is happening now, not what will happen next.
- **Industrial Drying Control Systems:** Commercial food processing plants use PLC-controlled dryers with PID feedback loops, but these systems cost tens of thousands of dollars and are designed for factory-scale operations, not open-air sun-drying used by small-scale processors.
- **Agricultural ML Platforms:** Companies like aWhere and CropX apply ML to agricultural sensor data for crop monitoring, but none target fish drying specifically, and their models are not transferable due to fundamentally different drying physics and environmental conditions

Technologies / Methods	Paper [3]	Paper [4]	Paper [5]	Paper [6]	Paper [7]	Proposed System
IoT Sensor Monitoring (Temperature, Humidity)	✗	✓	✓	✗	✗	✓
Real-Time Environmental Data Logging	✗	✓	✓	✗	✗	✓
Machine Learning Models	✗	✗	✗	✓	✓	✓
Drying Completion Time Prediction	✗	✗	✗	✓	✓	✓
Spoilage Risk Prediction	✗	✗	✗	✗	✗	✓
Time-Series Feature Engineering	✗	✗	✗	✓	✓	✓
Real-Time Risk Alerts	✗	✗	✓	✗	✗	✓
Mobile Dashboard Integration	✗	✗	✗	✗	✗	✓
Low-Cost Deployment for Small Processors	✗	✓	✓	✗	✗	✓
Integrated Predictive Decision Support	✗	✗	✗	✗	✗	✓

Table 1: Comparison of Existing Research Methods and the Proposed System

1.3 Research Gap

Despite the growing adoption of Internet of Things (IoT) technologies and machine learning techniques in food processing and agricultural drying, the literature review reveals several critical gaps in their application to dry fish processing, particularly in tropical developing countries such as Sri Lanka.

Gap 1: Monitoring Without Predictive Intelligence.

Existing IoT-based drying systems primarily focus on real-time environmental monitoring. These systems collect and display sensor readings such as temperature and humidity to allow operators to observe current drying conditions [4][5]. However, they do not provide predictive insights regarding the drying process. Specifically, current systems do not estimate the remaining drying time or predict potential spoilage risks based on evolving environmental conditions. As a result, processors must rely heavily on personal experience and manual judgment when deciding whether a batch will complete drying successfully or if environmental conditions may lead to spoilage. This lack of predictive capability limits the usefulness of IoT monitoring systems as decision-support tools.

Gap 2: Lack of Domain-Specific Feature Engineering for Fish Drying.

Previous studies on machine learning applications in agricultural drying demonstrate that predictive performance significantly improves when temporal features such as rolling averages, trend slopes, and rates of change are incorporated into the models [6][7]. However, these approaches have largely been applied to crops such as fruits and grains rather than fish drying environments. Important features specific to fish drying—such as humidity trend slopes, sudden humidity spikes caused by rainfall, weight loss rates during moisture evaporation, and cumulative exposure to high-humidity conditions—have not been systematically defined or incorporated into predictive models. The absence of domain-specific feature engineering limits the effectiveness of machine learning methods for fish drying prediction.

Gap 3: Limited Real-World Deployment of Machine Learning Models.

Many machine learning models developed for food drying prediction have been evaluated primarily in controlled laboratory environments. While these studies demonstrate promising predictive performance, they rarely translate into practical, deployable systems that can operate in real-world drying environments. Small-scale fish processors, particularly in developing regions, often lack access to advanced technological infrastructure and expertise. Consequently, there remains a gap between experimental research prototypes and field-ready solutions that can support everyday drying operations.

Gap 4: Absence of Integrated Prediction Frameworks.

Existing research typically focuses on a single predictive task, such as estimating drying kinetics or detecting spoilage conditions. Few studies attempt to integrate multiple predictive objectives within a unified system. In particular, there is limited work that simultaneously predicts drying completion time and spoilage risk levels using the same sensor data stream. An integrated predictive framework that combines both regression-based drying time estimation and classification-based spoilage risk prediction remains largely unexplored in fish drying applications.

To address these gaps, this research proposes an IoT-integrated predictive analytics framework for dry fish processing. The system combines real-time sensor monitoring with machine learning models capable of dual predictions: estimating drying completion time through regression models and predicting spoilage risk levels through classification models. A domain-specific feature engineering pipeline is introduced to extract meaningful temporal patterns from time-series sensor data, including rolling averages, humidity trend slopes, weight loss rates, temperature fluctuations, cumulative exposure to risk conditions, and sudden environmental changes. The study further evaluates both baseline models (Linear Regression and Logistic Regression) and ensemble learning methods (Random Forest and Gradient Boosting) to determine the most effective approach for this application. Finally, the predictive models are deployed through a FastAPI-based prediction service integrated with a Flutter mobile dashboard, enabling processors to receive real-time predictions and alerts. This approach bridges the gap between theoretical research and practical deployment, providing a field-ready decision-support system for sustainable dry fish processing.

2. Objectives

2.1 Main Objective

To design and develop a machine learning-based predictive analytics module that analyzes IoT sensor time-series data. This module will estimate drying completion time and predict spoilage risk for dry fish processing in Sri Lanka. It will help processors make informed decisions that reduce post-harvest losses

2.2 Specific Objectives

1. To design and deploy an IoT sensor data collection pipeline using ESP32 microcontrollers with DHT22 (temperature/humidity) and HX711 load cell (weight) sensors to capture time-series drying environment data at regular intervals.
2. To engineer domain-specific features from raw sensor time-series data. This includes rolling averages, humidity trend slopes, weight loss rate, temperature variation, cumulative time above risk thresholds, day/night flags, and sudden humidity jump events, to improve model predictive performance
3. To develop a regression model (Random Forest Regressor and Gradient Boosting Regressor) to estimate the remaining drying time for active fish batches based on current and historical sensor readings and engineered features.
4. To develop a classification model (Random Forest Classifier and Gradient Boosting Classifier) to predict the spoilage risk level (low/medium/high) of active batches based on environmental conditions and drying progress trajectory.
5. To evaluate and compare model performance using MAE and RMSE for regression, and Accuracy, Precision, Recall, F1-score, and confusion matrices for classification, benchmarking ensemble methods against baseline models (Linear Regression, Logistic Regression).
6. To deploy the trained models as a RESTful prediction service using FastAPI, integrated with the platform's Flutter mobile dashboard and Node.js backend, enabling real-time drying progress estimates and spoilage alerts for processors in the field.

3. Methodology

3.1 Research Approach

This research follows the Design Science Research (DSR) methodology, which is the standard framework for creating and evaluating innovative IT artifacts that solve real-world problems. DSR is the appropriate choice because the research does not merely study an existing phenomenon but designs, builds, and evaluates a novel technical artifact the ML predictive analytics module against real constraints of Sri Lanka's dry fish processing industry.

3.2 Technical Methods

Data Collection

IoT sensors are deployed at fish drying sites to collect time-series data at regular intervals of 5–10 minutes, enabling continuous monitoring of environmental conditions and drying progress. Each recorded data point represents the state of the drying environment and the batch at a specific moment. The collected dataset includes temperature, relative humidity, batch weight, drying duration, and the final batch outcome label, which are used as key features for analyzing the drying process and training the predictive machine learning models.

Temperature (°C):

Temperature is measured using the DHT22 sensor, which captures the ambient temperature of the drying environment. This parameter is important because the drying rate of fish is highly influenced by surrounding temperature. Higher temperatures generally accelerate moisture evaporation, while lower temperatures may slow down the drying process.

Relative Humidity (%):

Relative humidity is also measured using the DHT22 sensor and represents the amount of moisture present in the surrounding air. Humidity levels significantly affect the efficiency of the drying process. High humidity can slow down evaporation and increase the risk of spoilage, whereas lower humidity supports faster and safer drying.

Batch Weight (grams):

Batch weight is measured using a load cell integrated with an HX711 amplifier. This measurement tracks the gradual reduction in weight as moisture evaporates from the fish during the drying process. Monitoring weight loss over time provides valuable insight into the drying progress and helps estimate when the fish reaches the desired dryness level.

Drying Duration (minutes):

Drying duration represents the elapsed time since the batch was placed for drying. This variable helps in understanding the temporal progression of the drying process and is used as an important feature for modeling drying behavior and predicting the remaining drying time.

Batch Outcome Label:

After the drying process is completed, each batch is assigned a final outcome label based on a quality assessment. The batch is categorized either as successfully dried or assigned a spoilage risk level (low, medium, or high). These labels are used as target variables in machine learning models for classification and predictive analysis of drying quality.

Feature Engineering

Raw sensor readings are transformed into a set of engineered features that capture temporal dynamics critical for prediction accuracy [10]:

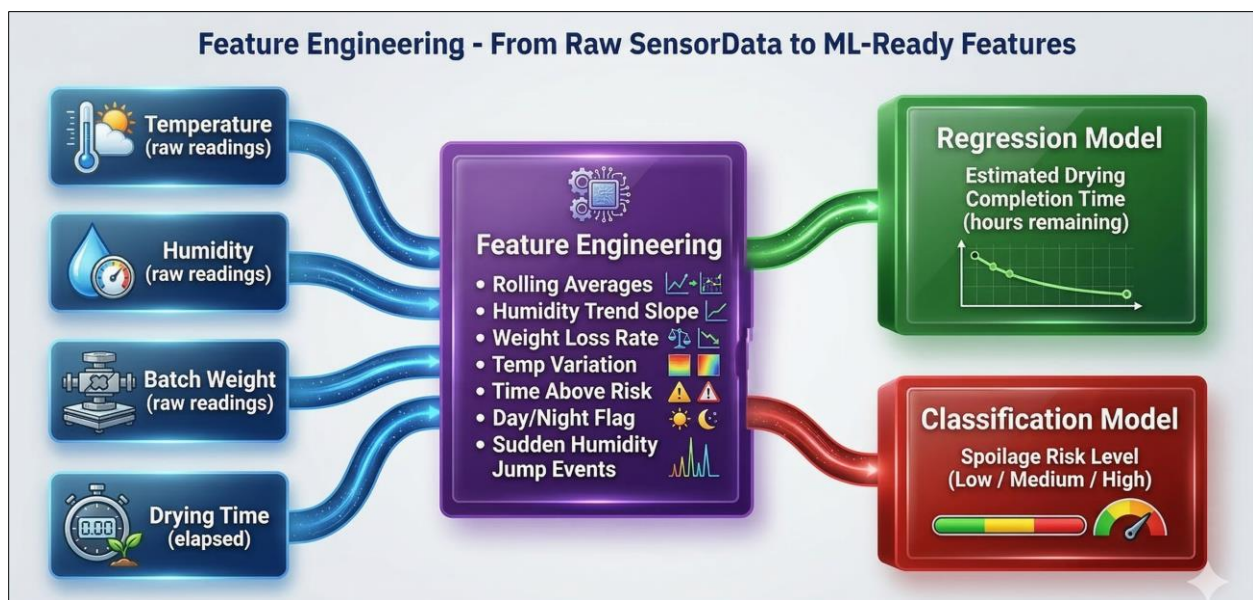


Figure 1: Feature Engineering Pipeline

Feature	Description	Rationale
Rolling Averages	Mean temperature, humidity, and weight over sliding windows (1h, 3h, 6h)	Smooths noise, captures sustained trends
Humidity Trend Slope	Linear regression slope of humidity over the last N readings	Detects rising or falling humidity trajectory
Weight Loss Rate	First derivative of weight over time (grams/hour)	Primary indicator of drying progress [11]
Temperature Variation	Standard deviation of temperature over sliding window	Captures day/night cycling and weather instability
Time Above Risk Threshold	Cumulative minutes where humidity > 75% or temp < 25°C	Proxy for microbial growth risk [9]
Day/Night Flag	Binary indicator based on time of day (6 AM–6 PM = day)	Captures diurnal drying rate differences
Sudden Humidity Jump	Binary flag when humidity increases > 15% in 30 minutes	Indicates rain onset or weather change event

Table 2: Sensor Data Features

Model Training and Evaluation

The predictive analytics module addresses two prediction tasks: drying completion time estimation and spoilage risk prediction. Separate machine learning pipelines are used for each task with appropriate algorithms and evaluation metrics.

Task 1 – Drying Completion Time Estimation (Regression)

This task predicts the remaining drying completion time based on IoT sensor data such as temperature, humidity, and weight. A Linear Regression model is used as a baseline, while Random Forest Regressor and Gradient Boosting Regressor are implemented as primary models to capture more complex relationships in the data.

Model performance is evaluated using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The dataset is divided using an 80/20 train–test split, and 5-fold cross-validation is applied to ensure model robustness.

Task 2 – Spoilage Risk Prediction (Classification)

This task classifies the spoilage risk level of fish into categories such as low, medium, and high based on environmental and drying conditions.

A Logistic Regression model is used as the baseline classifier, while Random Forest Classifier and Gradient Boosting Classifier are used as primary models to improve predictive performance.

The models are evaluated using Accuracy, Precision, Recall, and F1-score, along with a Confusion Matrix to assess classification performance. A stratified 80/20 train–test split is applied to maintain balanced class distribution.

Prediction Task	Model Type	Algorithms Used	Evaluation Metrics	Validation Method
Drying Completion Time Prediction	Regression	Linear Regression (Baseline), Random Forest Regressor, Gradient Boosting Regressor	MAE, RMSE	80/20 Train-Test Split + 5-Fold Cross Validation
Spoilage Risk Prediction	Classification	Logistic Regression (Baseline), Random Forest Classifier, Gradient Boosting Classifier	Accuracy, Precision, Recall, F1-score, Confusion Matrix	Stratified 80/20 Train-Test Split

Table 3: Machine Learning Model Training Setup

3.3 Tools and Platforms

Component	Technology	Justification
Data Processing	Python, Pandas, NumPy	Industry-standard for data manipulation and numerical computation
Machine Learning	Scikit-learn	Mature library with Random Forest, Gradient Boosting, and evaluation tools
Development Environment	Google Colab / Jupyter Notebook	Free GPU access; collaborative and reproducible
Prediction API	FastAPI (Python)	High-performance async framework; auto-generated API docs
Database	MongoDB	Flexible schema ideal for time-series sensor data with varying batch sizes
Mobile Frontend	Flutter	Cross-platform mobile dashboard (shared with team)
Backend	Node.js	Event-driven backend for real-time data handling (shared with team)
IoT Hardware	ESP32 + DHT22 + HX711 Load Cell	Low-cost, Wi-Fi-enabled microcontroller with proven sensor compatibility
Version Control	Git / GitHub	Source code management and collaboration

Table 4: Tools and Platforms

3.4 Data Collection and Ethical Considerations

Sensor data is collected from drying sites with the consent and cooperation of local fish processors. No personally identifiable information is collected from the sensors—only environmental readings (temperature, humidity, weight) and batch metadata (start time, fish type, batch size). Interview data from stakeholder consultations is anonymized before analysis. The project adheres to SLIIT’s ethical research guidelines, and informed consent is obtained from all interview participants.

3.5 Validation Strategy

The developed models are evaluated to measure their prediction accuracy and practical usefulness. For the drying completion time prediction, the predicted drying times are compared with the actual drying completion times recorded for test batches. The model performance is measured using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), with a target MAE of ≤ 60 minutes.

For the spoilage risk prediction, the predicted risk levels are compared with the actual outcomes of the drying batches. The models are evaluated using Accuracy, Precision, Recall, and F1-score, with particular attention to Recall for the high-risk class to ensure spoilage events are not missed.

In addition, the performance of baseline models (Linear/Logistic Regression) is compared with ensemble models (Random Forest and Gradient Boosting) to identify improvements in prediction accuracy. Finally, selected processors test the mobile dashboard with live predictions and provide feedback on the clarity and usefulness of the system.

Evaluation Area		Method	Target / Purpose
Drying Time Prediction	Time	MAE, RMSE	MAE ≤ 60 minutes
Spoilage Prediction	Risk	Accuracy, Precision, Recall, F1-score	High Recall for high-risk class
Model Comparison		Compare baseline and ensemble models	Show performance improvement
User Feedback		Dashboard testing with processors	Evaluate usability and usefulness

Table 5: Model Validation Metrics and Evaluation Criteria

4. High-Level System Architecture

4.1 System Overview

The proposed system operates as part of a broader Smart IoT-Driven Digital Platform for Sustainable Dry Fish Processing, which consists of four interconnected functional modules. This proposal specifically focuses on the Machine Learning–Based Predictive Analytics Module, which integrates with the other platform components to enable intelligent monitoring and decision support.

The Predictive Drying & Spoilage Analytics Module interacts with the following system components:

- **IoT Monitoring Module:** Responsible for collecting real-time environmental data such as temperature, humidity, and weight using ESP32-based sensors and transmitting the data to the backend system.
- **Computer Vision Quality Inspection Module:** Performs automated dry fish quality grading and defect detection using image analysis techniques.
- **Sustainability and Traceability Module:** Manages batch records, processing details, and waste tracking to support sustainable processing practices.

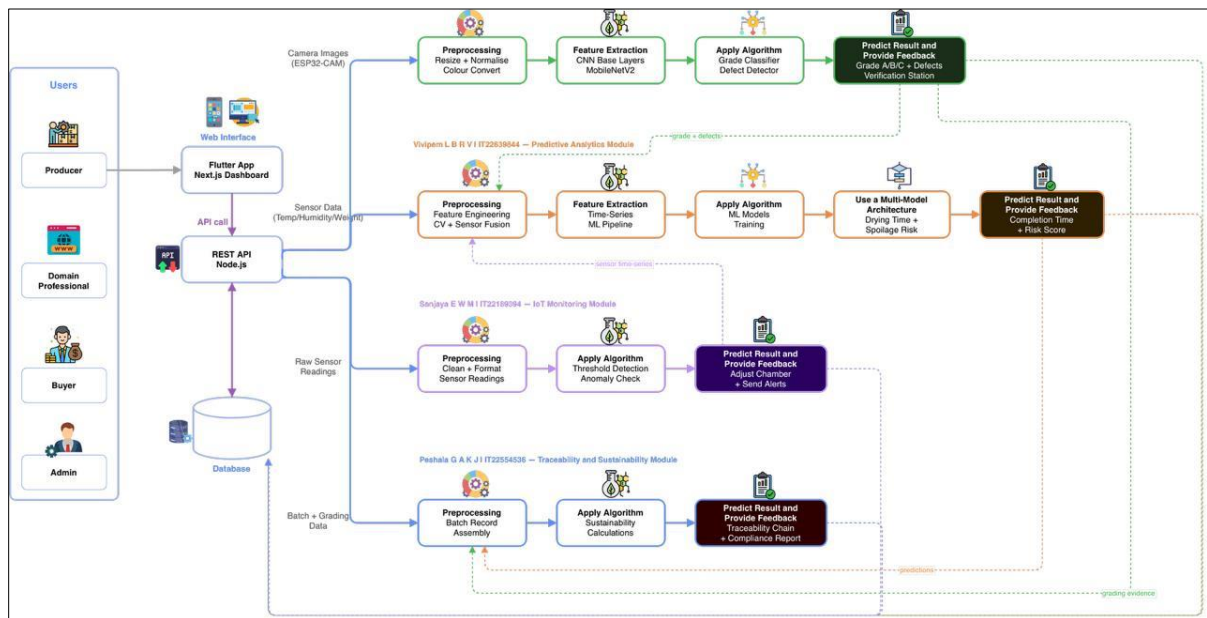


Figure 3: High Level System Architecture Diagram

The system follows a modular architecture, where each component operates as an independent service while communicating through a shared backend infrastructure. A Node.js backend server acts as the central communication layer, while MongoDB is used as the primary database for storing sensor data, processing records, and prediction outputs.

The IoT Monitoring Module continuously collects environmental sensor readings and stores the time-series data in the database. The Predictive Analytics Module retrieves this data, performs preprocessing and feature engineering, and applies trained machine learning models to estimate drying completion time and spoilage risk levels. The generated predictions are then written back to the database.

The mobile dashboard application, developed using Flutter, communicates with the backend to retrieve sensor data and prediction results. This enables processors to monitor drying progress in real time and receive early warnings about potential spoilage risks, supporting more informed decision-making during the drying process.

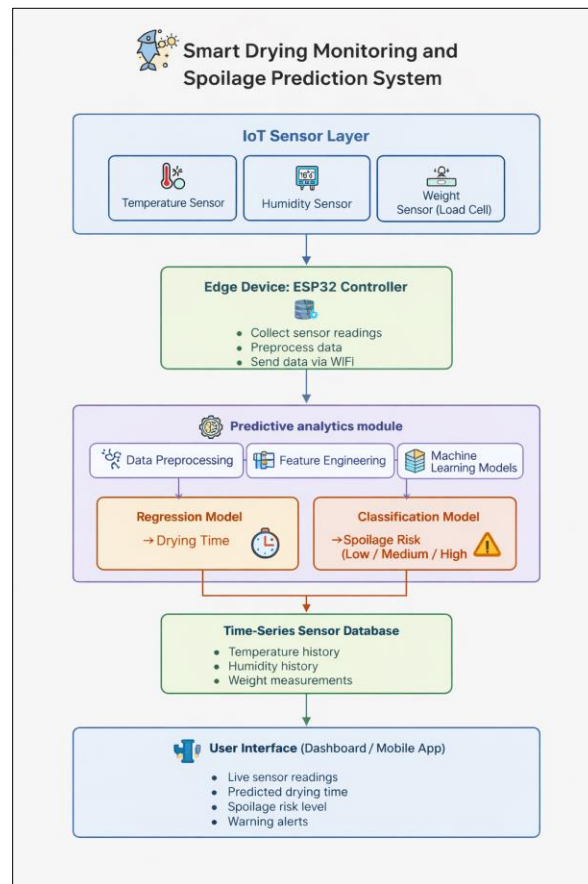


Figure 4: Individual System Diagram

Figure 3 illustrates the architecture of the Smart Drying Monitoring and Spoilage Prediction System, which integrates IoT sensing, machine learning analytics, and a user-facing dashboard to support dry fish processing operations.

At the IoT Sensor Layer, environmental and process data are collected using three sensors: a temperature sensor, a humidity sensor, and a weight sensor (load cell). These sensors continuously monitor the drying environment and the weight of the fish during the drying process.

The collected sensor readings are transmitted to the Edge Device, which is an ESP32 microcontroller. The ESP32 is responsible for gathering sensor measurements, performing basic preprocessing, and sending the data to the backend system through a Wi-Fi connection.

The data is then processed by the Predictive Analytics Module, which performs several machine learning operations. First, data preprocessing cleans and prepares the incoming sensor data. Next, feature engineering extracts meaningful patterns from the time-series readings. These features are then used by machine learning models to generate predictions. Two types of models are used: a regression model to estimate the remaining drying time and a classification model to predict the spoilage risk level (low, medium, or high).

All sensor readings and prediction results are stored in a time-series sensor database, which maintains historical records of temperature, humidity, and weight measurements. This database supports both model inference and system monitoring.

Finally, the processed information is delivered to the User Interface, which can be accessed through a dashboard or mobile application. The interface displays live sensor readings, predicted drying completion time, spoilage risk levels, and warning alerts, allowing processors to monitor the drying process and take timely actions to prevent spoilage or over-drying.

4.2 ML Prediction Pipeline

The core prediction workflow of the Predictive Drying & Spoilage Analytics Module follows a six-step pipeline:

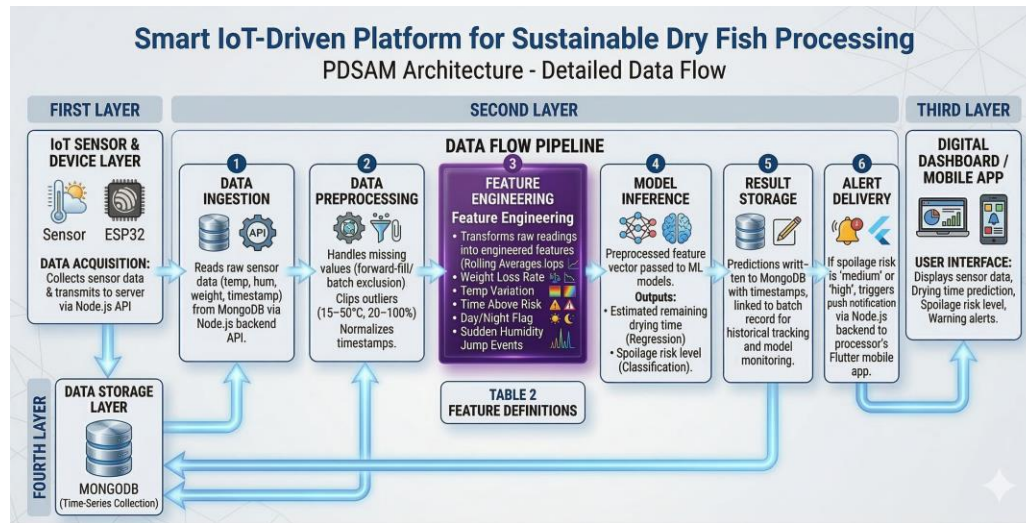


Figure 5: ML Prediction Pipeline (6 Steps)

Data Ingestion: Raw sensor data (temperature, humidity, weight, timestamp) is read from the MongoDB time-series collection via the Node.js backend API.

Data Preprocessing: Missing values are handled (forward-fill for brief gaps, batch exclusion for extended outages), outliers are clipped using domain-specific bounds (e.g., temperature: 15–50°C, humidity: 20–100%), and timestamps are normalized.

Feature Engineering: Raw readings are transformed into the engineered feature set (rolling averages, humidity slopes, weight loss rate, etc.) as defined in Table 2.

Model Inference: The preprocessed feature vector is passed to the trained ML models—the regression model outputs estimated remaining drying time (hours), and the classification model outputs spoilage risk level (low/medium/high) with confidence scores.

Result Storage: Predictions are written back to MongoDB with timestamps, linked to the specific batch record for historical tracking and model performance monitoring.

Alert Delivery: If the spoilage risk classification is "medium" or "high," the system triggers a push notification via the Node.js backend to the processor's Flutter mobile app.

4.3 Data Flow and Integration

The end-to-end data flow from sensor to processor follows this sequence:

- ESP32 sensors at the drying site read temperature (DHT22), humidity (DHT22), and weight (HX711 load cell) at 5–10 minute intervals and transmit readings via Wi-Fi to the Node.js backend.
- The Node.js backend validates and stores raw sensor data in the MongoDB time-series collection, making it available to all platform modules.
- The FastAPI prediction service periodically (or on-demand) fetches the latest sensor readings for each active batch, computes engineered features, and runs inference.
- Predictions (remaining drying time, spoilage risk level) are stored in MongoDB and made available through the backend API.
- The Flutter mobile dashboard polls the API and displays the current status of each batch: estimated completion time, spoilage risk indicator (color-coded: green/yellow/red), and historical drying curve.
- If spoilage risk reaches "medium" or "high," a push notification is sent to alert the processor.

4.4 Scalability and Security

Scalability Considerations:

The system is designed to support future expansion as more drying sites are added. The FastAPI prediction service is stateless, allowing it to be horizontally scaled by deploying multiple instances when needed. MongoDB time-series collections efficiently handle large volumes of continuous sensor data generated at short intervals. In addition, machine learning models can be periodically retrained using accumulated data through scheduled batch processes without interrupting the live prediction service

Security Considerations:

Security is ensured through several measures. Data transmitted from ESP32 sensors to the backend is protected using HTTPS/TLS encryption. The FastAPI prediction service is secured with API key authentication to prevent unauthorized access. The system does not collect any

personally identifiable information, storing only environmental sensor data and batch metadata. Access to MongoDB is restricted to authorized services using role-based database accounts.

5. User Requirements

5.1 Stakeholder Identification

Stakeholder Type	Stakeholder	Description / Role
Primary Stakeholders	Fish Processors (~50,000 coastal operators)	Primary end-users who receive drying predictions and spoilage alerts on their mobile devices.
Primary Stakeholders	Fish Traders and Wholesalers	Use batch quality data and drying completion estimates for procurement planning.
Primary Stakeholders	Mr. Dinuth Rashmika Jayasekara (External Supervisor)	Dry fish business owner providing domain expertise and validation.
Secondary Stakeholders	Project Team Members	IoT module provides sensor data; Computer Vision (CV) module consumes quality labels.
Secondary Stakeholders	System Administrators	Manage model deployment, retraining schedules, and alert thresholds.
Secondary Stakeholders	NARA / Fisheries Authorities	Potential consumers of aggregated analytics for fisheries management and policy planning.

Table 6: Stakeholder Identification

5.2 Functional Requirements

ID	Feature	Description
FR-01	Sensor Data Ingestion	The system shall ingest time-series sensor data (temperature, humidity, weight) from ESP32/DHT22/HX711 sensors at 5–10 minute intervals.
FR-02	Feature Engineering	The system shall automatically compute engineered features (rolling averages, trend slopes, weight loss rate, etc.) from raw sensor streams.
FR-03	Drying Time Prediction	The system shall estimate the remaining drying time (in hours) for each active batch using trained regression models.
FR-04	Spoilage Risk Classification	The system shall classify the current spoilage risk of each active batch as low, medium, or high using trained classification models.
FR-05	Real-Time Alerts	The system shall send push notifications to the processor's mobile device when spoilage risk reaches medium or high.
FR-06	Batch Dashboard	The Flutter mobile app shall display each active batch's estimated completion time, spoilage risk level, and historical drying curve.
FR-07	Prediction API	The FastAPI service shall expose RESTful endpoints for on-demand prediction requests from the mobile app and backend.
FR-08	Model Retraining	The system shall support periodic model retraining with accumulated batch data to improve prediction accuracy over time.

Table 7: Functional Requirements

5.3 Non-Functional Requirements

ID	Category	Description
NFR-01	Performance	The prediction API shall respond within 2 seconds for a single batch prediction request.
NFR-02	Accuracy	The drying time regression model shall achieve $MAE \leq 60$ minutes; the spoilage classifier shall achieve $F1\text{-score} \geq 0.85$.
NFR-03	Availability	The prediction service shall maintain 99% uptime during active drying seasons.
NFR-04	Usability	The mobile dashboard shall use simple color-coded indicators (green/yellow/red) understandable to users with limited technical literacy.
NFR-05	Scalability	The system shall support monitoring of at least 50 concurrent active batches without performance degradation.
NFR-06	Data Integrity	Sensor data shall be validated and missing readings handled gracefully (forward-fill) without producing erroneous predictions.

Table 8: Non-functional Requirements

6. Commercialization Plan

6.1 Target Market and Customer Persona

Sri Lanka’s dry fish processing industry represents a significant but underserved market for technology-driven quality improvement. The primary target segments are:

Small-Scale Coastal Processors (~50,000 operators): These are the primary users who process fish in open-air sun-drying yards along the coastal belt. They typically process 50–200 kg per batch, earn LKR 30,000–60,000 per month, and own basic smartphones. Their key pain point is unpredictable batch losses from weather changes and inconsistent drying outcomes.

Fish Traders and Wholesalers (~5,000): These intermediaries buy dried fish from processors and sell to retail markets and exporters. They need objective quality assurance and predictable supply timelines. Currently, they rely on visual inspection and personal trust relationships.

Dry Fish Exporters (~200 licensed exporters): Exporters to South Asian and Middle Eastern markets face increasing quality compliance requirements from importing countries. Objective sensor-verified quality data and traceability records give them a competitive advantage in international markets.

6.2 Value Proposition

The proposed predictive analytics system provides practical benefits to multiple stakeholders within the dry fish processing value chain. By combining IoT-based monitoring with machine learning predictions, the system improves decision-making, product quality, and operational transparency.

For Processors

The system helps processors minimize production losses by providing early spoilage warnings and accurate drying completion estimates. With better visibility into drying conditions, processors can determine the optimal time to stop drying, avoiding both over-drying and product spoilage. This is expected to reduce batch losses by approximately 15–25%, improving overall processing efficiency and profitability.

For Traders

Traders benefit from objective, sensor-verified quality information for each processed batch before purchasing. Access to reliable drying and spoilage indicators reduces uncertainty in quality assessment, minimizes disputes between processors and buyers, and supports fair and data-driven pricing decisions.

For Exporters

Exporters gain improved traceability and quality documentation through sensor-generated batch records. These records can support quality verification and compliance with international food safety and export regulations, helping exporters build trust with importing markets and potentially obtain higher-value market opportunities.

6.3 Revenue Model

The proposed platform follows a multi-stage revenue model to support gradual adoption and long-term sustainability.

Phase 1 – IoT Sensor Kit Sales

Pre-configured IoT sensor kits, consisting of an ESP32 microcontroller with temperature, humidity, and weight sensors, will be sold to processors at an estimated price of LKR 3,000–5,000 per kit. The pricing includes a small markup over the hardware cost and helps generate initial revenue while expanding the user base.

Phase 2 – Freemium SaaS Model

A freemium software model will be introduced for the platform. Basic monitoring and prediction features will be provided free of charge for up to five active drying batches. Advanced features such as unlimited batch monitoring, historical analytics, export-quality reports, and priority alert notifications will be available through a monthly subscription of LKR 500–1,000.

Phase 3 – B2B Analytics Services

Aggregated analytics derived from the collected sensor data, such as regional production volumes, quality trends, and seasonal drying patterns, will be offered as a data analytics service for traders and exporters. These insights can support supply planning and quality verification, with an estimated subscription cost of LKR 5,000–15,000 per month.

Phase 3 – Government and Institutional Partnerships

The platform can also be licensed to fisheries authorities and research organizations for large-scale monitoring of dry fish production. Such partnerships can support national programs aimed at improving product quality, reducing post-harvest losses, and enhancing traceability within the fisheries sector.

6.4 Pricing Strategy

Phase 1 (Pilot, Months 1–6): Free access to the full platform for pilot users at 3–5 drying sites. Sensor kits provided at cost. Goal: validate the technology and collect training data.

Phase 2 (Launch, Months 7–18): Free basic tier to drive adoption; premium subscription for power users. Sensor kits sold at 30–40% margin.

Phase 3 (Scale, Year 2+): Full pricing in effect. B2B data services launched. Target: 2,000 active processor subscriptions within 2 years.

6.5 Competitive Advantage

The proposed system introduces one of the first machine learning–based solutions for dry fish drying prediction in Sri Lanka, bringing data-driven decision support to a traditionally experience-based process. The platform is designed to be highly affordable, using low-cost IoT hardware such as ESP32, DHT22 sensors, and HX711 load cells, with an estimated total cost of around LKR 3,000, making it accessible to small-scale processors.

The machine learning models are specifically trained for local conditions, considering Sri Lankan fish species, tropical climate factors, and traditional open-air drying practices. The system is integrated with a simple mobile dashboard, allowing processors to monitor drying progress and receive spoilage alerts through smartphones without requiring technical expertise. Additionally, collaboration with an industry-based supervisor from a dry fish business provides real-world validation and practical insights for the solution.

6.6 Intellectual Property Considerations

- The trained ML models and engineered feature set for fish drying prediction constitute novel intellectual property eligible for software copyright protection.
- The integrated IoT-to-prediction pipeline architecture may be eligible for utility model registration under Sri Lanka’s IP framework.
- The accumulated sensor dataset from Sri Lankan drying sites is a unique asset with significant value for future research and commercial applications in the fisheries sector.

7. Budget and Justification

The following table outlines the estimated costs for developing and deploying the Machine Learning–Based Predictive Analytics Module over an 8-month research period. All costs are estimated in Sri Lankan Rupees (LKR).

Category	Item Description	Estimated Cost (LKR)
IoT Hardware	ESP32 dev boards (x3), DHT22 sensors (x3), HX711 + Load Cells (x3), breadboards, wiring, waterproof enclosures	10,000
Cloud / Compute Services	MongoDB Atlas (free tier + paid for production), Google Colab Pro for model training, server hosting for FastAPI service (8 months)	10,000
Software & Tools	Python, Scikit-learn, Pandas, NumPy, FastAPI, Flutter SDK, Node.js – all free/open-source	0
Communication & Travel	Travel to coastal drying sites for sensor deployment, field testing, and stakeholder interviews	10,000
Printing & Documentation	Report printing, poster for presentation	5,000
Contingency (10%)	Buffer for unforeseen hardware replacements or additional cloud costs	20000
TOTAL		55,000 LKR

Table 9: Budget Estimation

IoT Hardware Justification:

Three complete sensor kits are required: one for the primary drying site deployment, one for a secondary site to collect diverse environmental data, and one spare for hardware failures in the harsh coastal environment (salt air, high humidity, direct sunlight). The ESP32 + DHT22 + HX711 combination was chosen as the most cost-effective option that provides all three required measurements (temperature, humidity, weight) with Wi-Fi connectivity.

Cloud / Compute Justification:

Google Colab Pro provides GPU-accelerated model training without requiring expensive local hardware. MongoDB Atlas handles time-series data storage with automatic scaling. FastAPI hosting requires a lightweight cloud instance (approx. LKR 2,000–3,000/month) to serve predictions to the mobile app.

8. Work Breakdown Structure

#	Task	Description	Duration
1	Literature Review & Problem Analysis	Review IoT/ML for fish drying; field visits to drying sites; stakeholder interviews; FAO/NARA data review	4 weeks
2	IoT Sensor Deployment	Set up ESP32 + DHT22 + HX711 at drying sites; configure data transmission to MongoDB; validate sensor accuracy	3 weeks
3	Data Collection	Collect time-series sensor data across multiple drying cycles; label batch outcomes (success/spoilage)	6 weeks
4	Data Preprocessing & Feature Engineering	Handle missing data; remove outliers; implement feature extraction pipeline (rolling averages, slopes, etc.)	3 weeks
5	Model Training – Regression	Train Linear Regression, Random Forest Regressor, Gradient Boosting Regressor; hyperparameter tuning; cross-validation	3 weeks
6	Model Training – Classification	Train Logistic Regression, Random Forest Classifier, Gradient Boosting Classifier; evaluate with F1-score, confusion matrix	3 weeks
7	FastAPI Service Development	Build prediction API endpoints; integrate trained models; connect to MongoDB; add authentication	3 weeks

8	Mobile Dashboard Integration	Integrate prediction API with Flutter dashboard; implement batch status display and alert notifications	3 weeks
9	Testing & Evaluation	Model accuracy validation; API performance testing; user acceptance testing with processors	3 weeks
10	Documentation & Submission	Final report; poster preparation; code documentation; model handoff	3 weeks

Table 10: Work Breakdown Structure

9. Gantt Chart

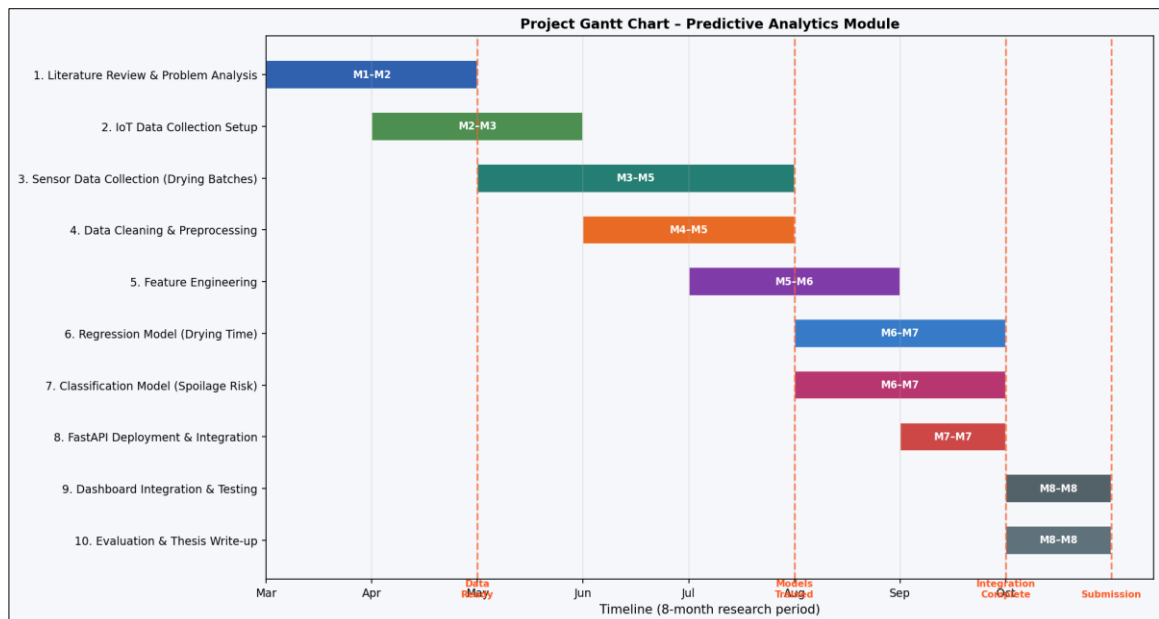


Figure 6: Project Gantt Chart

The Gantt chart above illustrates the 8-month project timeline across ten major work packages. Key milestones include: Sensor Deployment Complete (end of Month 2), Data Collection Complete (end of Month 4), Models Trained and Evaluated (end of Month 6), System Integration Complete (end of Month 7), and Final Submission (Month 8). Note that data collection (Task 3) overlaps with early preprocessing work, as initial batches can be processed while later batches are still being collected, maximizing the available training dataset.

References

- [1] Food and Agriculture Organization of the United Nations (FAO), "The State of World Fisheries and Aquaculture 2024: Blue Transformation in Action," FAO, Rome, 2024. [Online]. Available: <https://doi.org/10.4060/cd0683en> [1]
- [2] National Aquatic Resources Research and Development Agency (NARA), "Post-Harvest Losses in the Sri Lankan Dry Fish Industry: Status and Recommendations," NARA Technical Report, Colombo, Sri Lanka, 2022.
- [3] R. P. Jayasinghe, S. S. K. Haputhanthri, and A. D. Ranasinghe, "Assessment of post-harvest losses in traditional dry fish processing in Southern Sri Lanka," *Sri Lanka J. Aquatic Sciences*, vol. 26, no. 1, pp. 45–58, 2021.
- [4] A. Fudholi, K. Sopian, M. H. Ruslan, and M. Y. Othman, "IoT-based monitoring system for solar drying of agricultural products," *J. Advanced Research in Fluid Mechanics and Thermal Sciences*, vol. 56, no. 1, pp. 1–12, 2019.
- [5] K. Srinivasan and R. Kumar, "Smart fish drying system using IoT sensors and threshold-based alerts," in *Proc. IEEE Int. Conf. on IoT and Its Applications (ICIOT)*, Bangalore, India, 2020, pp. 112–117.
- [6] A. Motevali, S. Minaei, M. H. Khoshtaghaza, and H. Amirnejat, "Comparison of machine learning approaches for prediction of drying kinetics of agricultural products," *Drying Technology*, vol. 37, no. 16, pp. 2048–2060, 2019. [Online]. Available: <https://doi.org/10.1080/07373937.2018.1561005>
- [7] M. Kaveh, R. Amiri Chayjan, and E. Golpour, "Gradient boosting regression for modeling drying behavior of fruits under varying environmental conditions," *Computers and Electronics in Agriculture*, vol. 168, art. no. 105108, 2020.
- [8] O. A. Odeyemi, N. Alegbeleye, M. Strateva, and D. Stratev, "Understanding spoilage microbiology of modified-atmosphere-packaged seafood: A review," *Int. J. Food Microbiology*, vol. 285, pp. 18–24, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijfoodmicro.2018.06.016>
- [9] P. S. Taoukis and T. P. Labuza, "Time-Temperature Indicators (TTIs)," in *Novel Food Processing Technologies*, G. V. Barbosa-Canovas et al., Eds. Boca Raton, FL: CRC Press, 2005, ch. 5, pp. 103–126.
- [10] Y. Bouzembrak and H. J. P. Marvin, "Prediction of food safety hazards using machine learning on food safety time-series data," *Food Control*, vol. 61, pp. 35–45, 2016. [Online]. Available: <https://doi.org/10.1016/j.foodcont.2015.09.015>

- [11] V. Singh, A. K. Gupta, and S. P. Singh, "Weight loss kinetics during fish drying: Mathematical modeling and experimental validation," J. Food Engineering, vol. 246, pp. 112–120, 2019.
- [12] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," J. Management Information Systems, vol. 24, no. 3, pp. 45–77, 2007. [Online]. Available: <https://doi.org/10.2753/MIS0742-1222240302>
- [13] Department of Fisheries and Aquatic Resources, Ministry of Fisheries, "Fisheries Statistics 2023," Sri Lanka, 2023. [Online]. Available: <https://www.fisheriesdept.gov.lk>
- [14] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available: <https://doi.org/10.1023/A:1010933404324>

Appendices

Appendix A – Survey Instruments (Problem Validation)

Target Audience: Dry Fish Processors and Fish Traders

Sample Questions for Dry Fish Processors:

1. How many years have you been involved in dry fish processing?
2. On average, how many batches of fish do you dry per week, and what is the typical batch size (kg)?
3. How do you currently determine when a batch of fish is adequately dried? (Visual inspection / Touch / Time-based / Other)
4. In the past 6 months, approximately how many batches have you lost or had quality issues with due to over-drying, under-drying, or spoilage?
5. What weather conditions cause the most problems for your drying process? (Unexpected rain / High humidity / Temperature drops / Other)
6. If you had a mobile app that predicted when your fish batch would be ready and warned you about spoilage risk, would you use it? (Yes / No / Maybe)
7. Do you currently use a smartphone? If yes, which type? (Android / iPhone / Basic smartphone)

Sample Questions for Fish Traders:

8. How do you currently assess the quality of dried fish when purchasing from processors?

9. How often do you reject batches due to quality issues (under-dried, spoiled, contaminated)?
10. Would you be willing to pay a premium for dried fish that comes with sensor-verified quality data? (Yes / No / Depends on price)
11. What is the biggest challenge you face in sourcing consistently high-quality dry fish?

Appendix B – Risk Analysis

Risk Type	Description	Impact	Likelihood	Mitigation
Technical	Sensor hardware failure due to harsh coastal environment (salt air, moisture)	High	Medium	Deploy waterproof enclosures; maintain spare sensor kits; implement data gap handling in preprocessing
Data Quality	Insufficient training data due to limited drying cycles during collection period	High	Medium	Supplement with synthetic data augmentation; use cross-validation with smaller datasets; extend collection period if needed
Model	ML models overfit to specific drying site conditions and fail to generalize	Medium	Medium	Collect data from multiple sites; use regularization; validate on held-out site data
Integration	Latency in prediction delivery due to network connectivity issues at coastal sites	Medium	High	Implement local caching on mobile app; batch prediction requests; offline-capable dashboard
User Adoption	Processors resist using mobile app due to low digital literacy	Medium	Medium	Simplify UI to minimal color-coded indicators; provide Sinhala/Tamil language support; conduct hands-on training

Table 11: Risk Analysis