Advance Food Grading System
(AFG)

By

Ahmer Husnain

Alina Haider Syed

Muhammad Haseeb

Muhammad Saad Abdullah

Supervised by:

Assoc Prof. Dr. Muhammad Imran

Submitted to the faculty of Department of Electrical Engineering,
Military College of Signals, National University of Sciences and Technology, Islamabad,
in partial fulfillment for the requirements of B.E Degree in Electrical Engineering.

June 2023

In the name of ALLAH, the Most benevolent, the Most Courteous

CERTIFICATE OF CORRECTNESS AND APPROVAL

This is to officially state that the thesis work contained in this report

"Advance Food Grading System"

is carried out by

Ahmer Husnain

Alina Haider Syed

Muhammad Haseeb

Muhammad Saad Abdullah

under my supervision and that in my judgment, it is fully ample, in scope and excellence, for the degree of Bachelor of Electrical Engineering in Military College of Signals, National University of Sciences and Technology (NUST), Islamabad.

Approved by

Supervisor
Assoc Prof. Dr. Muhammad Imran
Department of EE, MCS

Date: ___________

DECLARATION OF ORIGINALITY

We hereby declare that no portion of work presented in this thesis has been submitted in support of another award or qualification in either this institute or anywhere else.

ACKNOWLEDGEMENTS

Allah Subhan'Wa'Tala is the sole guidance in all domains.

Our parents, colleagues and most of all supervisor, **Dr. Muhammad Imran** without your guidance.

The group members, who through all adversities worked steadfastly.

Plagiarism Certificate (Turnitin Report)

This thesis has _6_ _ similarity index. Turnitin report endorsed by Supervisor is attached.

Ahmer Husnain

00000306530

Alina Haider Syed

00000294027

Muhammad Haseeb

00000285557

Muhammad Saad Abdullah

00000296795

Signature of Supervisor

ABSTRACT

Near-infrared (NIR) spectroscopy is growing as a valuable non-destructive analytical technique, acquiring prominence in a variety of application areas. This thesis describes the creation of a low-cost, portable NIR-based device for classifying fruits and grading milk. Using NIR sensors AS7263, the device measures the absorption of light by fruit samples at six distinct wavelengths. The data gathered by the aforementioned sensors is processed by machine learning algorithms in order to classify produce and grade milk. The device includes a web application that enables real-time viewing of classification results, thereby facilitating the making of informed decisions. The results of principal component analysis (PCA) indicate that the proposed device can effectively differentiate between various fruits and accurately detect variable milk-water percentages. Two distinct R-based machine learning models demonstrate that the classif.ranger algorithm can effectively classify fruits based on their spectral reflectance values and milk samples based on their spectral reflectance data with an accuracy of 0.964% and 96.36%, respectively. The analysis verifies that machine learning algorithms can be utilized effectively in the agriculture and dairy industries for quality control and food safety applications. The proposed device provides a cost-effective alternative to the costly spectrometers presently on the market, which can be prohibitive for researchers and producers.

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Overview

Near-infrared (NIR) spectroscopy has become a valuable nondestructive analytical technique, which has gained remarkable importance in several fields of application. It has proven to be a tool of choice due to its primary advantages in practical roles such as its applicability to a wide variety of samples and rapid noninvasive analysis. These benefits are well-matched with autonomous, portable spectrometers that are capable of on-site analysis. In the past decade, the emergence of NIR spectrometers has led to a significant leap in the practical applications of this technique. NIR spectroscopy extracts information from the sample through molecular vibrational excitations, similar to IR and Raman techniques. However, it differs from these techniques in that only overtones and combination transitions can be observed in the NIR spectral region (typically defined as 12,500-4000 cm-1 or 800-2500 nm). It is noteworthy that the principles of NIR spectroscopy and its various applications have been extensively researched and documented [7].

1.2 Problem Statement:

The agriculture and dairy industries are crucial for the economy of many countries, including Pakistan. However, the quality of milk and fruits can be compromised due to adulteration, which can have severe health implications for consumers. Therefore, there is a need for accurate and efficient techniques to classify fruits and grade milk to ensure quality control.

The field of non-destructive food analysis has seen a growing interest in the development of software tools and data analysis methods that enhance the performance and reliability of NIR spectroscopy measurements. NIR spectroscopy has gained popularity due to its various advantages, such as rapid analysis and applicability to a wide range of samples. However, the cost of equipment remains a significant challenge for food processing industries and researchers who require frequent analysis. [4]

The agriculture and dairy industries are crucial for the economy of many countries, including Pakistan. However, the quality of milk and fruits can be compromised due to adulteration, which can have severe health implications for consumers. Therefore, there is a need for accurate and efficient techniques to classify fruits and grade milk to ensure quality control.

Near-infrared (NIR) spectroscopy has shown promise as a non-destructive and rapid technique for fruit classification and milk grading [8]. However, there are challenges associated with developing reliable and accurate models due to variations in spectral data caused by environmental factors or changes in sample composition.

1.3 Objectives:

1.3.1 General Objective:

The general objective of this project is to develop a reliable and accurate method for fruit classification and milk grading using NIR spectroscopy and machine learning algorithms.

1.3.2 Academic Objectives

To review the existing literature on NIR-based techniques for fruit classification and milk grading, including a discussion of the advantages and limitations of different types of NIR sensors and machine learning algorithms used in similar studies.

- ✔ To collect and analyze spectral data from different fruit and milk samples using a portable NIR spectrometer.
- ✔ To develop and compare machine learning models, such as decision trees, support vector machines, and neural networks, for accurately classifying different fruit types and grading milk based on its quality parameters.
- ✔ To evaluate the performance of the developed models by comparing their accuracy, precision, and recall metrics with the results obtained using traditional analytical methods.

- ✔ To identify the challenges associated with developing accurate and reliable models for fruit classification and milk grading using NIR spectroscopy, such as variations in spectral data due to environmental factors or changes in the composition of the samples.

- ✔ To propose recommendations for improving the accuracy and reliability of the developed models and suggest future research directions for enhancing the application of NIR spectroscopy and machine learning algorithms in the agriculture and dairy industries.

1.4 Scope:

The scope of our project is to the development and evaluation of NIR-based techniques for fruit classification and milk grading using machine learning algorithms. Specifically, we aim to collect spectral data from various fruits and milk samples using a NIR spectrometer, preprocess the data, and train machine learning models for accurate classification and grading.

1.5 Deliverables

The main deliverable of the project is a working prototype of the NIR-based fruit classification and milk grading system. This system will comprise the necessary hardware and software components required for accurate and efficient classification and grading.

The hardware components of the system will include an AS7263, a sample holder, and ESP-32 to control the spectrometer and process the spectral data. The NIR spectrometer will be responsible for measuring the NIR spectra of the fruit or milk samples, while the sample holder will securely hold the samples in place during measurements. The computer or microcontroller will be responsible for controlling the spectrometer, processing the spectral data, and running the machine learning algorithms for classification and grading.

The software components of the system will include data processing and analysis tools, as well as machine learning algorithms for classification and grading. The data processing and analysis tools will be used to preprocess the spectral data, such as baseline correction and smoothing, to remove any noise or artifacts in the data. The machine learning algorithms will be trained on the preprocessed spectral data to classify the fruit samples into different classes and grade the milk samples based on their fat and protein content.

Once the system is fully functional and validated, it can be used in various agricultural and dairy industries for accurate and efficient fruit classification

1.6 Relevant Sustainable Development Goals

The Sustainable Development Goals (SDGs) aim to ensure a sustainable and equitable future for all. Several SDGs are addressed by the initiative, including:

1. ***Life on land:*** This research helps to the sustainable management and protection of land resources by enabling precise fruit categorization and milk grading, which can help improve agricultural production and reduce food waste.

2. ***Decent work and economic growth:*** By providing a low-cost alternative to expensive spectrometers, the initiative encourages economic growth and innovation in the agriculture and dairy industries, resulting in the creation of new employment possibilities.

The project addresses the need for low-cost, dependable equipment for fruit categorization and milk grading, which are critical for maintaining quality standards and combating food adulteration. Accurate categorization and grading are also required for fair pricing and benefiting small-scale farmers' livelihoods. The project's emphasis on increasing agricultural output and minimizing food waste coincides with SDGs 1 and 2: No Poverty and Zero Hunger, respectively. SDG 8: Decent Work and Economic Growth is aligned with the creation of a low-cost alternative to costly spectrometers, which fosters economic growth and innovation in the agriculture and dairy industries. In addition, the initiative promotes sustainable land management practises, which are in line with SDG 15: Life on Land.

.

1.7 Structure of Thesis

Chapter 2: Literature review, background, and analysis study

Chapter 3: Design and development of the project

Chapter 4: Detailed evaluation and analysis of the code

Chapter 5: Results and discussion of the project

Chapter 6: Conclusion of the project

Chapter 7: Future work needed for commercialization of the project

Chapter 2: Literature Review

It has been demonstrated that NIR sensors are useful for figuring out the makeup of food items like fruits and milk. Milk grading and fruit categorization investigations have used both conventional and miniature NIR spectrometers. For instance, Zeb et al. (2019) classified various kinds of fruits based on their short-wave NIR spectral characteristics using a miniature NIR spectrometer. [2] Barthès (2019) examined the performance of a conventional and a miniaturized NIR spectrometer for characterizing soil nitrogen and carbon in a different investigation. Studies employing NIR-based methods for classifying fruits and milk have made substantial use of machine learning algorithms. [3] One of the common techniques used for this is neural networks, followed by decision trees and support vector machines. A support vector machine was employed in the work by Zeb et al. (2019) to categorize fruits based on their NIR spectral properties. [2] Similar to this, Furhan Iqbal (2019) developed a strategy for identifying milk adulteration using decision trees.[1] Despite the benefits of NIR-based methods for classifying fruits and grading milk, creating accurate and dependable models is not without its difficulties. The variation in spectral data brought on by environmental conditions or changes in the makeup of the samples is one of the main obstacles. For instance, the makeup of fruit might change depending on the region where it is grown, when it is harvested, and how ripe it is. The breed of the animal, the time of lactation, and the animal's nutrition can all affect the content of the milk.

2.1 Industrial background

In today's era, Pakistan's food business is fragmented and inefficient, with a lack of standardization and openness on food quality and safety. Concerns concerning food safety and public health have been raised as a result of the use of toxic pesticides and chemicals in food production. Due to a lack of knowledge and confidence in the food sector, consumers have difficulties making educated food purchasing decisions.

To solve these difficulties, several technologies and techniques have been created. NIR spectroscopy has evolved as a helpful method for food quality evaluation in recent years, notably for milk grading and fruit classification. For this aim, NIR sensors have been employed in both conventional and tiny NIR spectrometers. Machine learning algorithms like as neural networks, decision trees, and support vector machines have been used extensively in studies to construct reliable categorization models for fruits and milk.

The unpredictability in spectral data produced by climatic conditions and changes in sample composition, on the other hand, offers a considerable difficulty for developing reliable models. Fruit and milk composition might vary based on growth circumstances, harvest time, and animal nutrition, making it challenging to construct standardized models.

To solve these issues, novel portable devices with cloud-based data processing and interactive web apps have been created to give customers accurate food quality assessments and standardized grading systems. These technologies have the potential to transform the Pakistani food business by enhancing food safety and quality control, lowering the use of toxic pesticides and chemicals, boosting transparency and confidence in the food sector, and decreasing food waste.

2.2 Existing solutions and their drawbacks

Portable spectrometers like the SCiO, Viavi systems, ABB, and ASD are examples of NIR spectroscopy systems that are now available for milk and food grading. However, these remedies have a few shortcomings:

- ***High Cost:*** The pricing of many of the portable spectrometers now available range from $1,299 to $35,000. For small-scale farmers and researchers who do not have the funds to buy such equipment, this high cost might be a considerable obstacle.
- ***Limited Availability:*** Small-scale farmers and researchers who need portable spectrometers are unable to purchase them due to the high cost of these instruments. In areas where agriculture plays a big role in the economy, this restricted availability can be a serious problem.
- ***Complexity:*** A lot of the current solutions are complicated and need skilled experts to use them. Small-scale farmers and researchers who might lack the technical know-how to use such equipment may find this complexity to be a hurdle.
- ***Limited Uses:*** Some of the current solutions were created for particular uses, such soil analysis or milk quality testing, and may not be appropriate for other uses. For scientists and farmers who want a more adaptable solution, this narrow range of applications might be a considerable disadvantage.
- ***Destructive Testing:*** Some current options call for destructive testing, which necessitates destroying the food or milk sample in order to gather the required information. For farmers and academics who need to retain their samples for future study, this might be a serious downside.

In general, the current approaches to NIR spectroscopy in food and milk grading have serious shortcomings, including high costs, a lack of availability, complexity, a lack of applicability, and damaging testing. These restrictions make it difficult for academics and small-scale farmers to use this technology and obstruct the growth of an effective and sustainable food business.

Chapter 3: Materials And Method

3.1.1 Fruit Sample Collection

The fruit samples used in this study were obtained from the local market in Rawalpindi, Pakistan. We collected 30 dates, 30 strawberries, 21 kiwis, and 20 tomatoes as samples. The fruit samples were stored in a cool and dry place to prevent spoilage before analysis. The samples were then scanned using our device from different angles, and a dataset was created based on the number of readings taken for each fruit type. Specifically, we collected 52 readings for dates, 171 readings for kiwis, 98 readings for strawberries, and 67 readings for tomatoes. These readings were then used to train our machine learning model for accurate fruit classification. We also collected data on the mean and standard deviation of the spectral features of the fruit samples using the NIR-based portable device. The summary statistics for the fruit absorption levels in different wavelengths show that there are significant differences between the fruits. For example, the mean absorption levels at 610 nm for Dates is 1364.560, which is significantly lower than that for Kiwi (3899.578), Strawberry (2453.616), and Tomato (4080.216). Similarly, the standard deviation of absorption levels at 680 nm for Dates is much higher (323.9251) than that for Kiwi (101.92712), Strawberry (151.92683), and Tomato (274.50065). These findings suggest that the absorption levels of fruits vary significantly depending on the type of fruit and the wavelength of light.

Overall, Kiwi appears to have the highest absorption levels across all wavelengths, while Dates have the lowest absorption levels. Strawberries and tomatoes fall somewhere in between. Table 1 shows the results of our analysis of the spectral features of the fruit samples.

Fruits	count	mean_610nm	sd_610nm	mean_680nm	sd_680nm	mean_730nm	sd_730nm	mean_760nm	sd_760nm	mean_810nm	sd_810nm	mean_860nm	sd_860nm
Date	49	1364.560	378.0668	312.5885	90.00516	323.9251	74.28473	124.4918	35.09842	93.23897	23.68461	36.70491	8.743567
Kiwi	118	3899.578	389.5491	918.0523	101.92712	920.5193	91.29533	416.1466	45.96755	272.63257	24.25019	113.72168	10.458047
Strawberry	99	2453.616	510.4787	970.2180	151.92683	694.3237	96.32917	316.0246	40.90671	199.53077	26.88162	94.66816	12.246370
Tomato	68	4080.216	482.8647	1236.5651	274.50065	772.3100	143.59998	313.5348	68.90002	229.71061	39.08766	109.85630	19.120282

Table 1: Mean and standard deviation of spectral features of fruit samples

In addition, we have included two images of the fruit samples used in our study. Figure 1.a shows a sample of strawberries, and Figure 1.b shows samples of dates, tomatoes, and kiwis. These images can help readers visualize the types of fruits used in our study.

Figure 1.a: strawberries sample. Figure 1.b: dates, kiwi, and tomatoes sample.

3.1.2 Milk Grading Sample Collection:

We collected samples of milk to investigate the effect of dilution on the spectral features of milk. The milk samples were obtained from a local dairy farm in Rawalpindi, Pakistan. We used pure milk and added water in 10% increments to create a series of dilutions ranging from 0% to 100%. For each dilution level, we collected 10 samples, resulting in a total of 100 samples. The samples were collected in sterile containers and stored in a refrigerator at 4°C until analysis. We used a

NIR-based portable device to measure the spectral features of the milk samples. The device was calibrated before use using a standard reference material.

After collecting the spectral data, we calculated the mean and standard deviation of the spectral features for each dilution level. The results are presented in Table 2. As the percentage of water added to the milk increased, we observed a significant decrease in the absorption levels at all wavelengths. For example, the mean absorption levels at 610 nm for pure milk (0% dilution) were 13260.264, which decreased to 3919.237 for the 100% dilution. Similarly, the standard deviation of absorption levels at 680 nm increased as the percentage of water increased.

Our findings suggest that the spectral features of milk are highly sensitive to dilution. This has important implications for the dairy industry, where milk is often adulterated with water to increase its volume. By measuring the spectral features of milk, it may be possible to detect such adulteration and prevent the sale of low-quality milk.

Milk	count	mean_610nm	sd_610nm	mean_680nm	sd_680nm	mean_730nm	sd_730nm	mean_760nm	sd_760nm	mean_810nm	sd_810nm	mean_860nm	sd_860nm
0%	10	13260.264	3.919116	2711.9479	1.0897930	1674.6590	0.7154455	643.36200	0.5095483	594.29027	0.3060136	288.98820	0.0000000
10%	10	11441.898	63.703114	2633.7491	17.3621465	1412.6989	54.9628317	620.16035	4.6307156	567.87092	4.0238755	280.06018	2.2967114
100%	10	3919.237	6.552613	260.3461	0.7122946	530.5965	1.3263900	84.39595	0.6524930	94.74193	0.7140686	36.41722	0.5537825
20%	10	10329.388	4.866703	2551.9620	0.8321952	1319.3381	0.4994531	593.67187	0.3056974	541.54838	0.4997698	274.77382	0.3715044
30%	10	12457.141	717.160802	2619.7131	17.7789500	1597.9358	73.6707482	619.09698	6.8388491	571.45159	7.4018188	280.76496	2.4765694
40%	10	11242.218	4.234438	2531.2779	1.1342563	1338.6250	0.4469352	586.80800	0.0000000	546.00002	0.4079760	269.95734	0.4953392
50%	10	12612.906	3.218824	2519.0363	0.8618122	1570.0654	0.5468652	590.96501	0.4669605	552.19352	0.4997181	271.48448	0.3715044
60%	10	12402.502	9.157154	2391.5545	1.2973346	1534.1412	1.4174607	660.27988	1.2057224	564.38702	1.0986460	260.08944	0.9906415
70%	10	11597.230	485.590388	2260.0622	15.3593750	1402.4195	84.5296235	523.10019	13.2665720	496.64512	7.6852687	245.40507	2.6239152
80%	10	9889.132	3.596365	1919.3012	1.1180442	1219.8311	0.6014897	452.43188	0.6445000	429.29031	0.6766470	212.51211	0.6668492
90%	10	8479.975	5.267918	1372.0147	0.5992191	1013.7180	0.8938705	315.05883	0.3056974	314.90322	0.4997698	154.71446	0.5674823

Table 2: Mean and standard deviation of spectral features of milk samples at different dilution levels

3.2 Design of handheld NIR-based device:

3.2.1 AS7263

In this study, we used a NIR sensor AS7263 from Spark Fun to measure the absorbance by fruits. It is based on Near-Infrared (NIR) technology. It is a high-precision device that is designed to measure the intensity of light across different wavelengths within the NIR range.

The AS7263 sensor consists of two key components: the sensing element and the control electronics. The sensing element consists of a set of 18 photodiodes that are arranged into three different groups, each with six photodiodes. Each of these groups is optimized to detect light within a specific wavelength range, and they cover the range of 600 nm to 870 nm. The AS7263 spectrometer detects wavelengths in the visible range at 610, 680, 730, 760, 810, and 860 nm of light, each with 20nm of full-width half-max detection.

The control electronics of the AS7263 are responsible for powering the sensing element and reading the output of each photodiode. They are also responsible for converting the raw data into calibrated spectral data that can be easily interpreted and analyzed. The default current and light intensity of the AS7263 spectral sensor is 12.5mA and 1.2mA, respectively. These values were maintained throughout the development and testing.

3.2.2 ESP32

We utilized the ESP32 microcontroller with Micro python to design a smart IoT device. The ESP32 received data from a sensor, interacted with the user via RGB LEDs, and sent data to the Firebase cloud. A Wi-Fi manager was also programmed to make the device independent and connect to any available Wi-Fi network. The device collected and transmitted data in real-time, allowing for easy monitoring and quick response to any abnormalities detected. The use of Micro python reduced development time and made it easier to program the device.

3.2.3 RGB LED

RGB LEDs are a type of light-emitting diode that is capable of producing a wide range of colors by combining different intensities of red, green, and blue light. They consist of three separate diodes arranged in close proximity to each other, and are commonly controlled using a microcontroller or driver circuit with pulse-width modulation (PWM) to adjust the voltage across each diode. They also provided visual feedback to the user about the state of the device. Specifically, we used three colors to indicate different states of the device: red, blue, and green. When the device was turned on but not connected to Wi-Fi, the LED would glow red, indicating that the device was not yet ready for use. Once the device was connected to Wi-Fi and ready to use, the LED would change to blue. Finally, when the device was actively sending data to the Firebase cloud, the LED would change to green, indicating that the device was functioning as intended.

The use of RGB LEDs allowed for easy and intuitive monitoring of the device's status, even from a distance. The LEDs were connected to the ESP32 microcontroller and were programmed to respond to the state of the device in real-time. By providing this visual feedback, we were able to improve the usability of the device and make it more accessible to users with varying levels of technical expertise.

The table below summarizes the different colors of the RGB LED and their corresponding states:

LED Color	Device State
Red	Device on but not connected to Wi-Fi
Blue	Device connected to Wi-Fi and ready to use
Green	Device sending data to Firebase cloud

Table 3: Summary of the RGB LED colors and their corresponding device states

3.2.4 Pushbutton

We utilized two pushbuttons in the design of our IoT device. The first pushbutton was used as an on/off button, allowing the user to turn the device on or off as needed. The second push button was used to change the state of the device. Specifically, this push button was used to initiate the process of taking readings from the sensor and sending the data to the Firebase cloud. The pushbutton was also used to put the device in standby mode when readings were not being taken, thereby conserving power and extending the device's battery life. The use of these push buttons allows for easy and intuitive control of the device, while also providing the flexibility to adjust the device's operation based on the user's needs. The push buttons were connected to the ESP32 microcontroller, which was programmed to respond to the user's inputs and change the state of the device accordingly.

3.2.5 Rechargeable batteries:

Rechargeable batteries have a similar structure to disposable batteries, with two electrodes separated by an electrolyte. The electrodes and electrolytes in rechargeable batteries are designed to be more stable and reversible, allowing the battery to withstand multiple cycles of charging and discharging. Rechargeable batteries can be recharged using a charger.

Chapter 4: Device Setup and Operation

4.1 Device Setup

4.1.1 Connecting to Wi-Fi

When the device is switched on, it initiates a search for nearby Wi-Fi networks. If no familiar network is detected, the device starts broadcasting its own Wi-Fi signals with the name "WIFI Manager." Users can connect to this signal by entering the password set by them. Once connected, the device launches a small server that can be accessed through the web browser of the connected device. This process is indicated by the red light being switched on, indicating that the device is not connected to the internet.

4.1.2 Launching Wi-Fi Manager Portal:

Figure 4 (see below) shows the screenshot of the portal. The user interface is designed to be user-friendly, allowing users to easily connect to the WIFI by selecting the available options and entering the password. The device connects to the Wi-Fi smoothly within a short time span of 3-10 seconds, provided that the password entered is correct. If the password entered is incorrect, the user is informed and asked to try again. If the password entered is correct, the device is connected to the internet. This process is indicated by the green light being switched on, indicating that the connection with the internet has been established.

4.1.3 Connecting to the Internet:

Once the device is connected to the internet, users can easily switch to the data-sending mode by pressing the push button, resulting in the blue light and the sensor being ready to take readings. Users can exit this mode by long pressing the push button for 3-5 seconds, which changes the light to green, indicating that the device is back in the simply connected state.

Figures:

Figure 4: Screenshot of the Wi-Fi Manager Portal

Figure 5: Method of taking a reading from the device

4.2 Device Operation

4.2.1 Taking Readings

Taking a reading from our device is a simple process. Figure 5 shows how users can take readings from our device by placing fruit in front of the sensor in the equatorial position. It is important to place the fruit correctly to get consistent readings with the minimum effect of the surroundings. The fruit should cover the full surface above the sensor so that no external rays can interfere with the process.

The device starts taking readings, and the sensor gives a bright flash indicating that a reading has been taken. The sensor continues to take readings at regular intervals unless we switch the mode back or the device gets disconnected from the internet. The readings taken are sent to the database from where they are sent to our web app. These readings can be used for fruit classification, Lacto Grading or creating a dataset by saving them in a CSV file.

4.2.2 Sending Readings to the Database

After the device is set up and connected to the internet, the device can start taking readings from the fruit or milk samples placed in front of the sensor, as described in Section 3.3.2. These readings are then automatically sent to the database via the internet. The database is responsible for storing the readings, which can then be accessed by the web app for further analysis and processing. The readings can also be used for creating datasets of absorption values for different objects, which can be saved in a CSV file format. This feature is particularly useful for researchers and food industries who want to gather and analyze large amounts of data quickly and easily.

4.2.3 Switching Modes

The device has two modes: the connected state and the data-sending mode. The connected state is when the device is connected to the internet and the green light is switched on, indicating the connection is established. In this mode, users can easily switch to the data-sending mode by pressing the push button again, which results in the blue light and the sensor being ready to take readings. The readings are automatically sent to the database, as described in Section 3.3.1.

To exit the data-sending mode, users can long press the push button for 3-5 seconds, which changes the light to green, indicating that the device is back in the connected state. In this state, the device is still connected to the internet, but it is not taking readings or sending data to the database. State of the device is informed via RGB light colors. Related modes of the device with RGB light are shown in Table 3.

Overall, the device is designed to be user-friendly and simple to operate, with the ability to connect to the internet easily and switch between modes effortlessly. The combination of the device and web app allows for real-time data analysis and processing, making it a valuable tool for researchers and food industries. The figures mentioned in this section, such as Figure 5 and Figure 6, provide a visual representation of the device and web app interfaces for readers to better understand the operation and features of the system.

4.3 Web App Interface and Features.

4.3.1 User Interface

The user interface of the developed web app is designed to be user-friendly and intuitive. The interface is divided into two main components, the frontend and the backend.

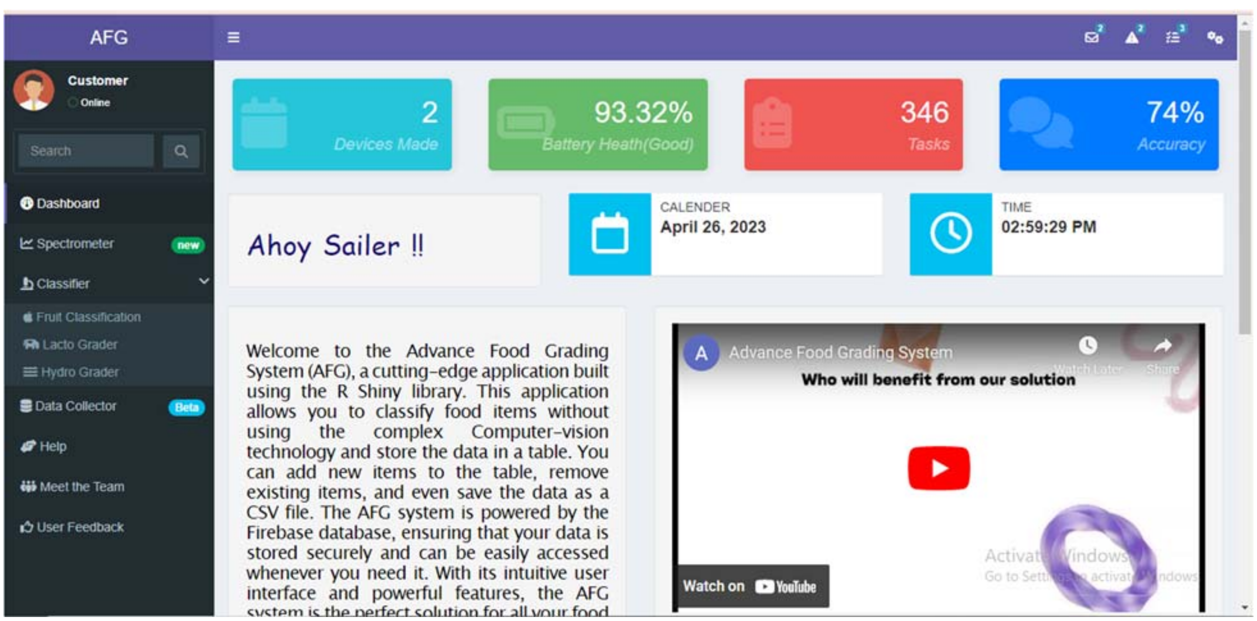


figure 6: web app interface-dashboard

Frontend

The frontend of the web app provides users with the tools to interact with the device and visualize the data collected. Figure 6 shows the user interface of the web app, which provides users with an absorption graph, fruit classifier, and Lacto Grader. The web app can also be used to create a dataset of absorption of different objects. figure 7 and figure 8 shows fruit classification and Lacto grader tab respectively. Also, the spectrum is shown in the spectrometer tab (see figure 9 below).

Backend

The backend of the web app is designed to process and display the data collected by the device. Figure 10 shows the backend dashboard of the web app, which is built using R shiny web app. The backend of the web app uses the Random Forest Classification Algorithm to classify fruits and determine water percentage in milk. The data is fetched from the database and processed to display the desired action to the user.

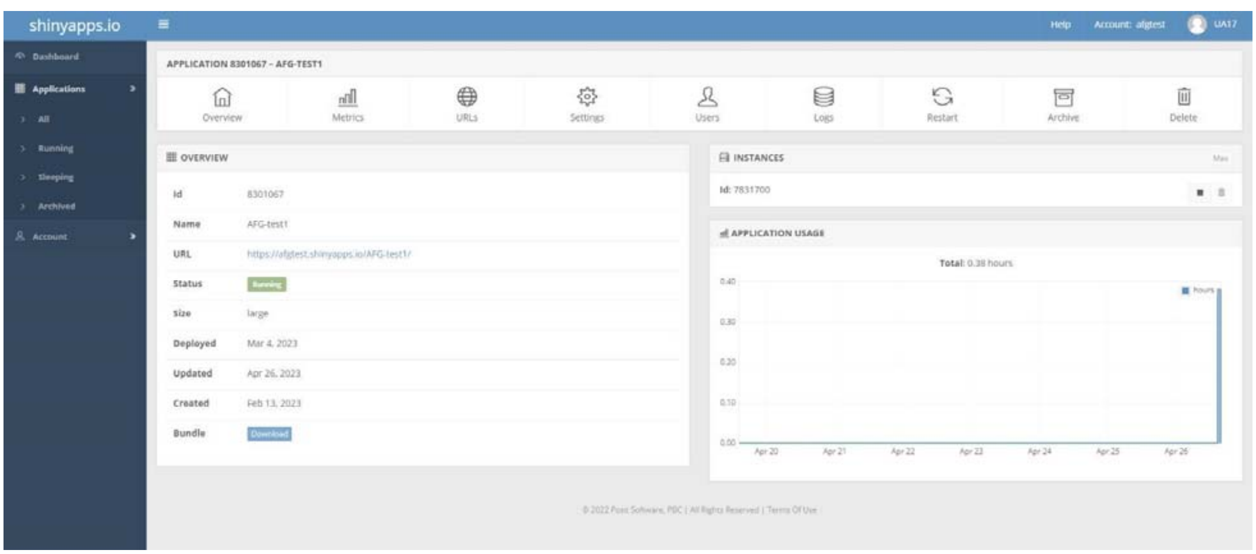


Figure 10: Backend of Web app

4.3.2 Absorption Graph

The absorption graph feature of the web app provides users with a real-time graph of the object scanned from the device, as shown in Figure 9. The graph updates in real-time and becomes static once the device stops sending readings. This feature of the web app acts as a spectrometer, providing users with great detail related to the absorption scheme for the object.

figure 9: spectrometer tab in web app

One of the key features of our web app is the absorption graph, which acts as a spectrometer and provides users with great detail related to the absorption scheme for the object scanned from the device. As shown in Figure 9, the absorption graph is updated in real-time as the readings are sent from the device to the web app. Once the device stops sending readings, the graph becomes static, providing users with a clear visual representation of the absorption spectrum.

The absorption graph is a valuable tool for researchers and food industry professionals who need to analyze the chemical properties of different substances. With our device and web app, users can easily obtain absorption spectra without the need for expensive equipment or specialized training.

4.3.3 Fruit Classification

Predicting Fruit Types

Our fruit classification feature, as shown in Figure 7, uses a machine learning algorithm to predict the type of fruit based on its spectral reflectance data. When a reading is taken through the sensor and sent to the web app, it is processed by the Random Forest Classification Algorithm implemented in the web app's backend. The predicted fruit type is then displayed to the user on the fruit classification tab.

The accuracy of the fruit classification feature is dependent on several factors, including the quality of the readings taken by the device, the accuracy of the algorithm used, and the size and diversity of the training dataset. In our study, we achieved an accuracy of 0.964 using the classif. ranger algorithm in R. However, the accuracy can be improved by increasing the number of training samples and optimizing the algorithm further.

Our fruit classification feature currently supports four fruits, as mentioned in the paper. However, more fruits can be easily added by updating the training dataset and retraining the algorithm.

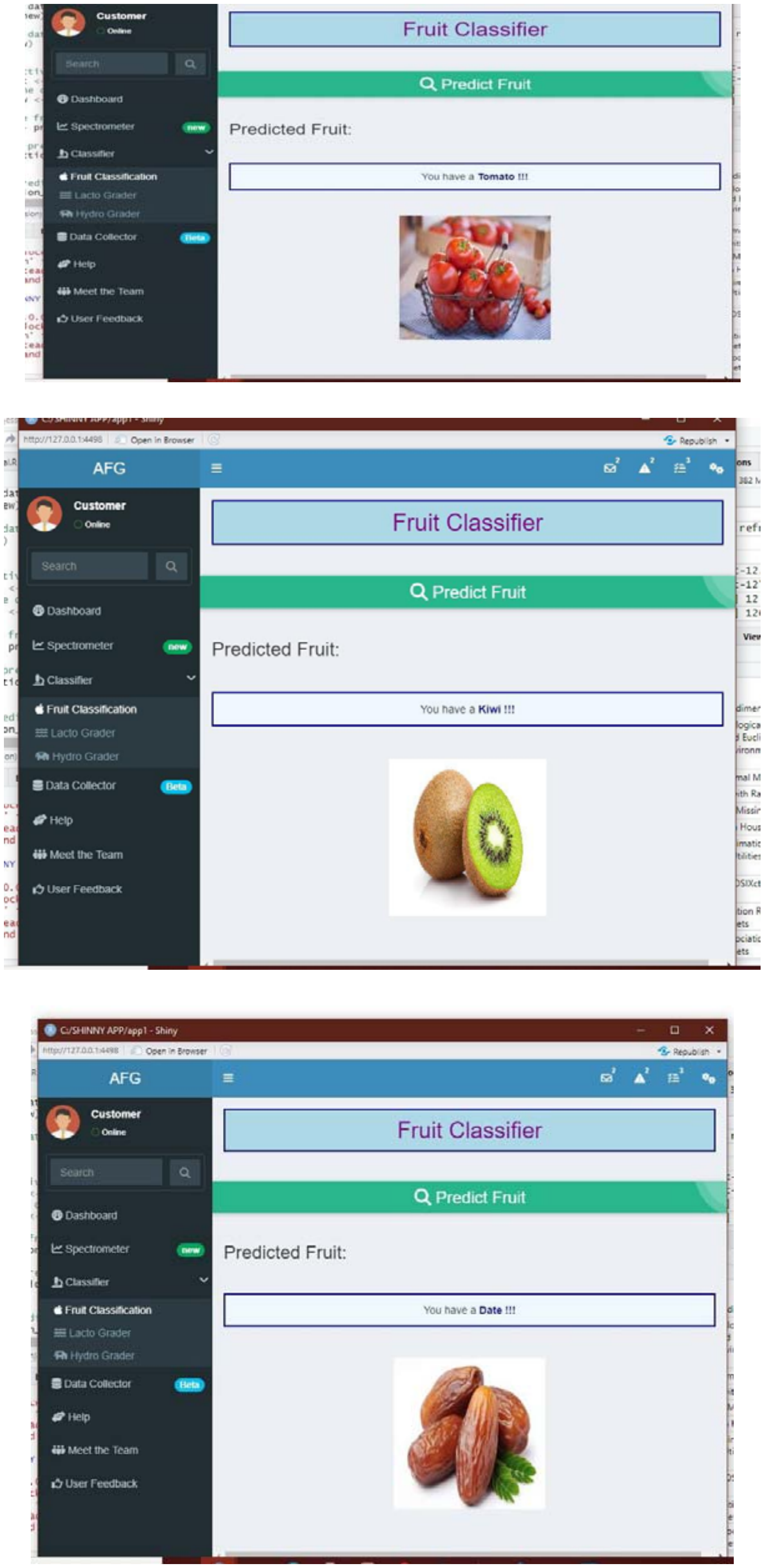


figure 7: Fruit Classification in Web Application

4.3.4 Lacto Grader

Lacto Grader is a unique feature of our portable NIR device that can detect the adulteration of water in milk. This feature is essential for ensuring the quality of milk and preventing health hazards associated with contaminated milk consumption. The Lacto Grader interface is designed to predict the water percentage in the milk solution by analyzing the spectral reflectance values obtained from the device.

Predicting Water Percentage

The user needs to place the milk sample in a small transparent container on top of the sensor and take a reading using the device. The Lacto Grader interface then displays the predicted water percentage in the milk solution. The accuracy of the prediction depends on the quality of the milk sample and the calibration of the device. Therefore, it is important to use the device according to the instructions provided in section 3.3.2.

The working principle of the Lacto Grader feature is based on the absorption of NIR radiation by water and other components in the milk sample. The presence of water in milk decreases the absorption of NIR radiation, leading to a decrease in spectral reflectance values. Thus, the lower the spectral reflectance value obtained from the milk sample, the higher the water percentage in the milk solution.

Testing Method

The Lacto Grader interface uses machine learning algorithms to predict the water percentage in the milk solution based on the spectral reflectance values obtained from the device. The Random Forest Classification algorithm is implemented in the backend of the web app to perform this task. The algorithm uses a set of decision trees to classify the milk sample into different water percentage categories. The accuracy of the prediction is calculated using the confusion matrix, which is a measure of the model's performance in predicting the correct water percentage category. Figure 8 shows the Lacto Grader interface of our web app. The interface displays the predicted water percentage in the milk solution and the accuracy of the prediction. The user can also view the absorption graph of the milk sample, which provides a detailed view of the spectral reflectance values obtained from the device.

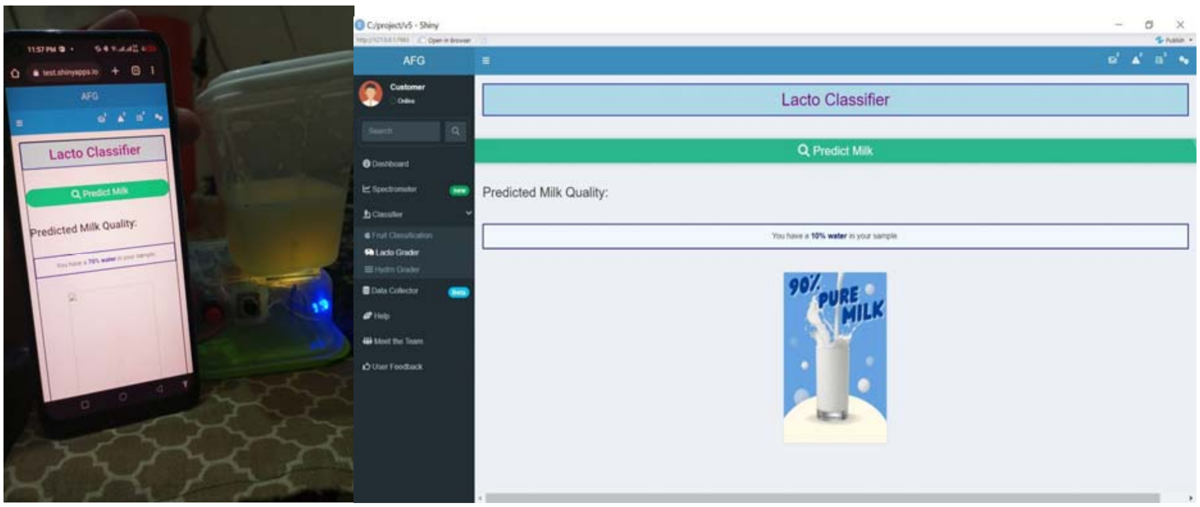


figure 8: Lacto-Grader tab in Web Application

On left we can see real time use of Mobile Application on Market Milk

On Right we can see Result of Lacto Grader on Desktop Web Application

4.3.5 Data Collection

The device not only allows the users to perform fruit classification and lacto-grading but also provides the functionality to create a dataset of the absorption spectra of various objects. This data can be used for further analysis and research purposes.

Creating a Dataset

The dataset creation feature of the device is easy to use and user-friendly. Users can create a dataset of any object by scanning it through the device and saving the data by clicking the "Add to Table" button in the data collector interface (see Fig. 11). The user can place the object in front of the sensor in the equatorial position and take readings as explained in section 3.3.2. Once the readings are taken, the user can save the data in the form of a CSV file by naming and saving it. The user can also remove any unnecessary data by using the "Remove from Table" button.

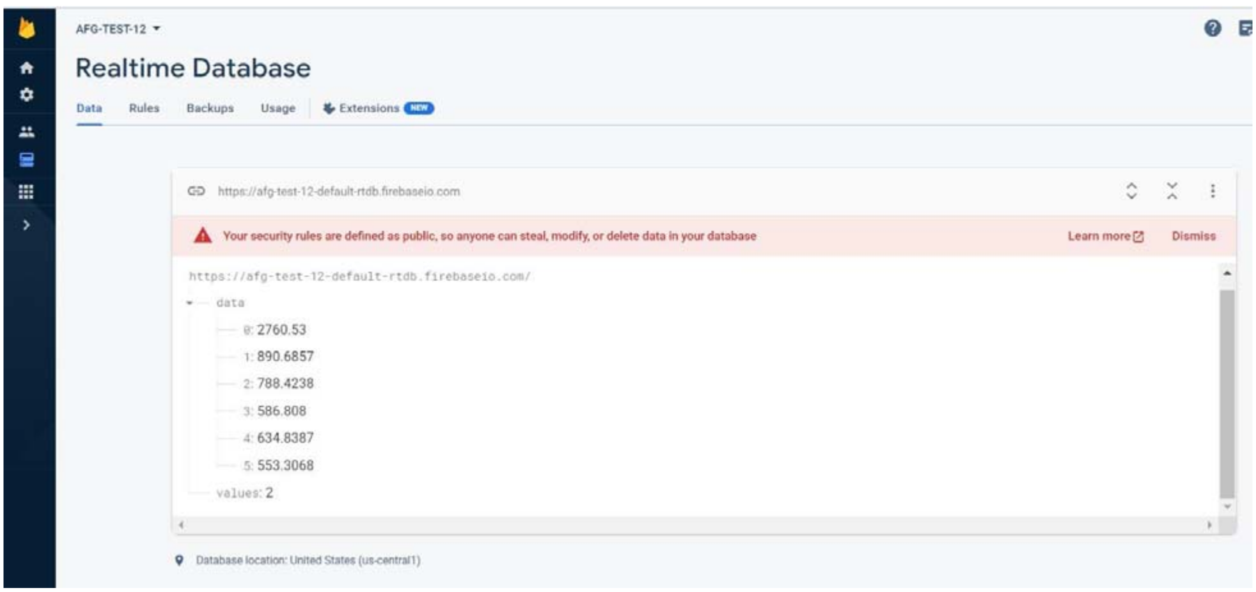


Figure 11(a): Realtime Database

Saving a Dataset

The created dataset can be used for further analysis and research. Users can save the dataset in the form of a CSV file, which can be easily imported into various data analysis software for further processing. The CSV file can be named as per the user's choice and saved in the desired location on the computer.

The data collection feature of the device is essential for creating a database of various objects' absorption spectra. This data can be used for developing new applications and algorithms for various industries such as agriculture, food processing, and quality control. The dataset creation feature provides researchers with an easy-to-use and efficient method of collecting data, making it easier to analyze and compare the spectral data of various objects.

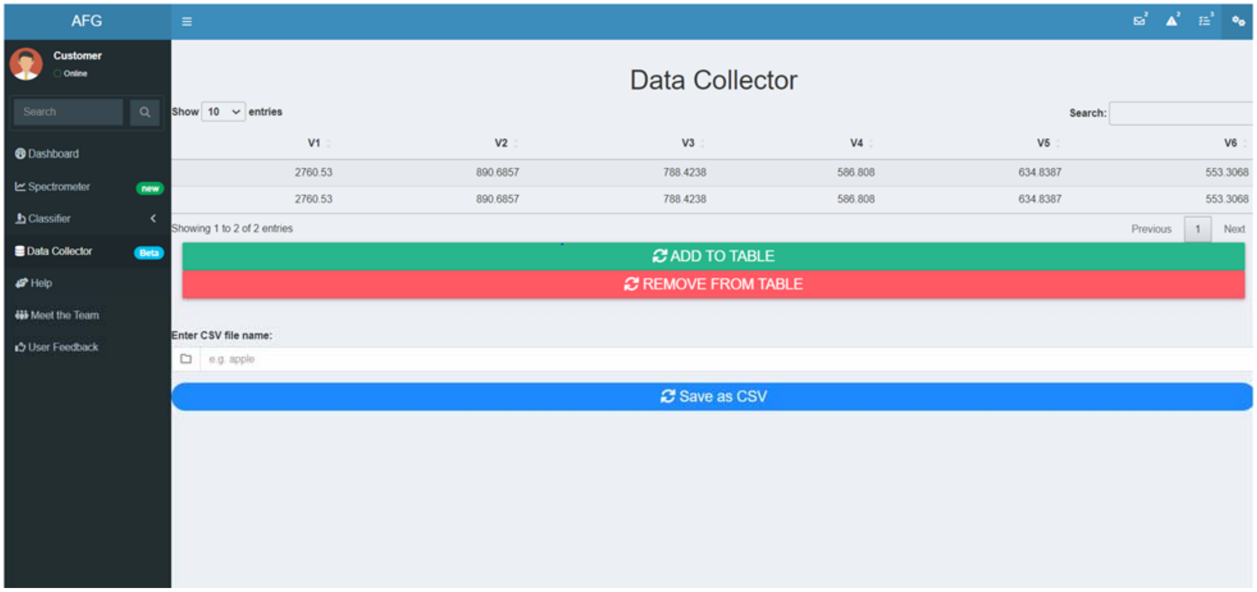


Figure 11(b): Data Collector Web Interface

Chapter 5: Result and Discussion

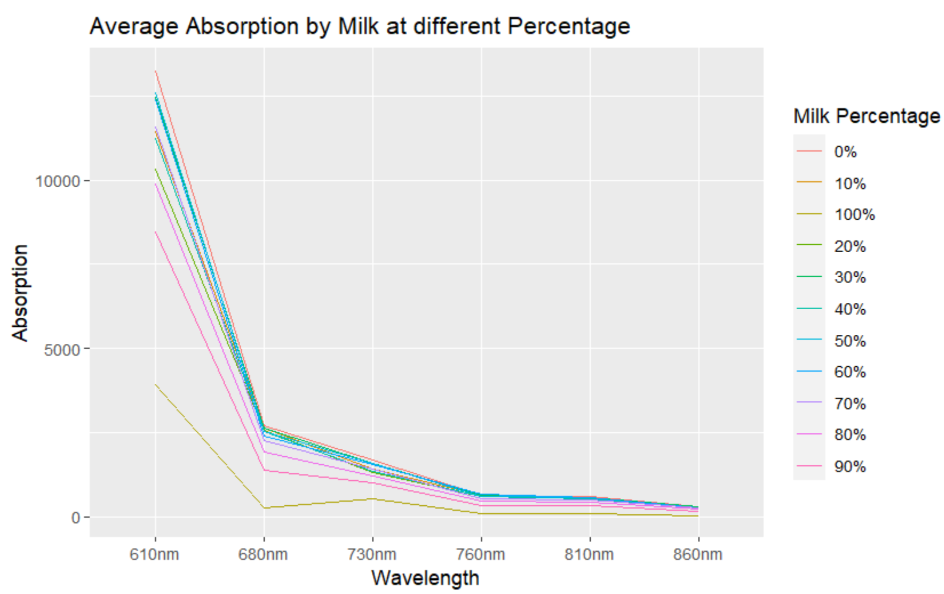
In this chapter, we will discuss the results and implications of our study on fruit classification and milk grading using NIR-based techniques and machine learning algorithms. We will begin by presenting the observations of the spectral data and the principal component analysis results. Then, we will discuss the classification models developed for fruit classification and milk grading, along with their performance measures. Finally, we will examine the implications of our findings for the agriculture and dairy industries, as well as the limitations of our study and areas for future research. Overall, our study provides important insights into the potential of NIR-based techniques and machine learning algorithms for improving the efficiency and accuracy of fruit classification and milk grading processes.

5.1 Observation of Spectra:

The collected data is analyzed by visualizing it. The mean absorption of a Milk-water solution at selected percentages to NIR wavelengths is presented in Fig. 2, while Fig. 1 represents a visual graph for the mean absorption of each fruit at selected wavelengths. The results are further validated through PCA analysis.

5.1.1 Principal Component Analysis

In our project, we performed PCA analysis on the collected data for both fruit classification and milk-water percentage detection. The scores plot for fruit classification is presented in Fig. 12(a), where the eleven investigated fruits formed well-defined clusters using six wavelengths from the AS7263 sensor. The score plot shows that our proposed device has the potential to distinguish fruits. For milk-water percentage detection, we performed PCA analysis on the mean absorption of Milk-water solution at selected percentages to NIR wavelengths. The results are presented in Fig. 12(b), which shows the distinct clusters formed by different milk-water solutions, indicating that our proposed device can accurately detect different milk-water percentages. The PCA analysis results demonstrate the effectiveness of our proposed device for both fruit classification and milk-water percentage detection. [5]



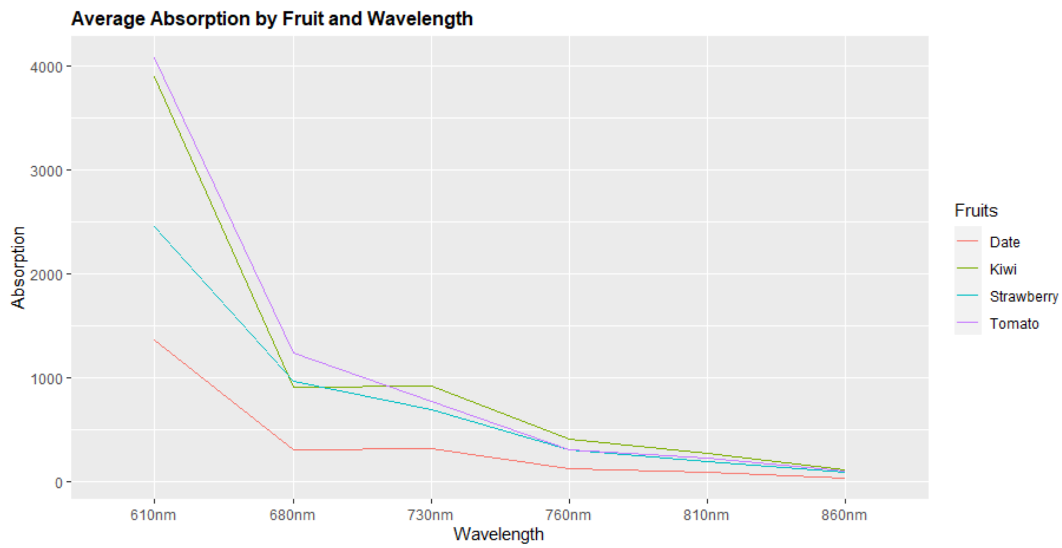
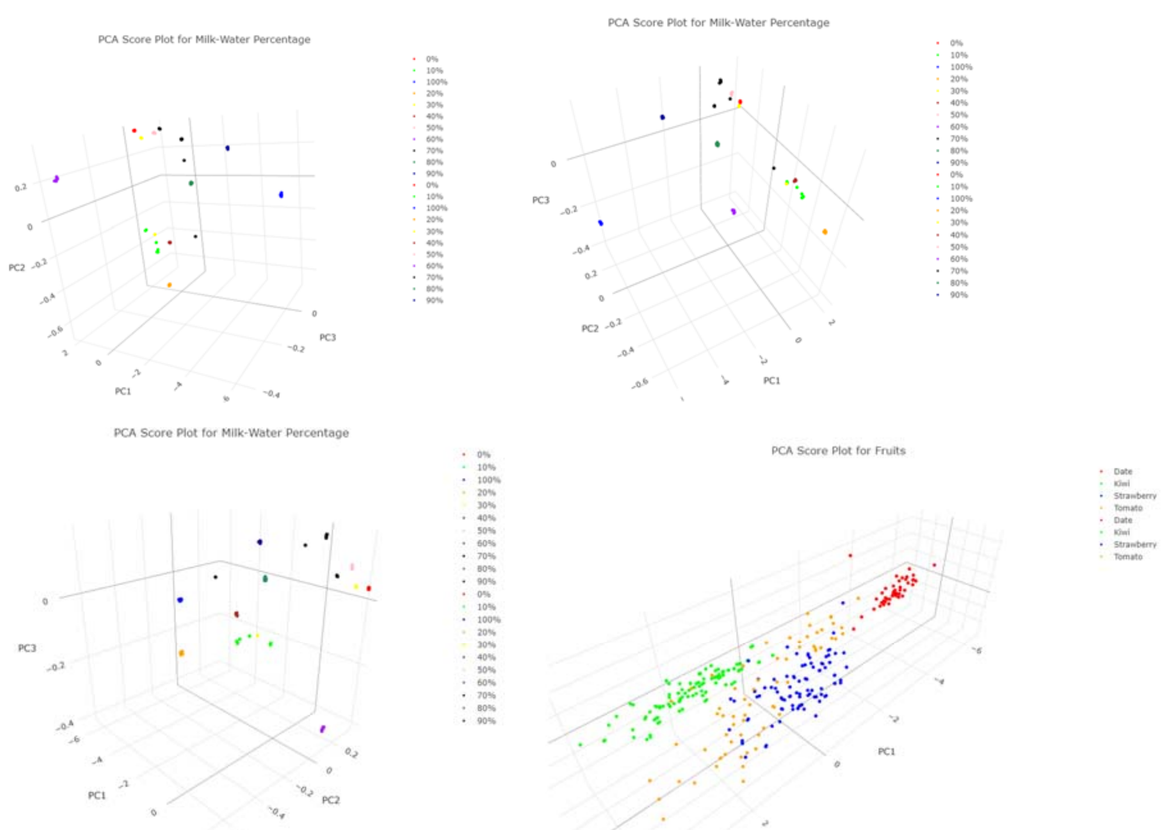


Figure 12 (a): Line Charts showing Absorption for samples at various Wavelengths



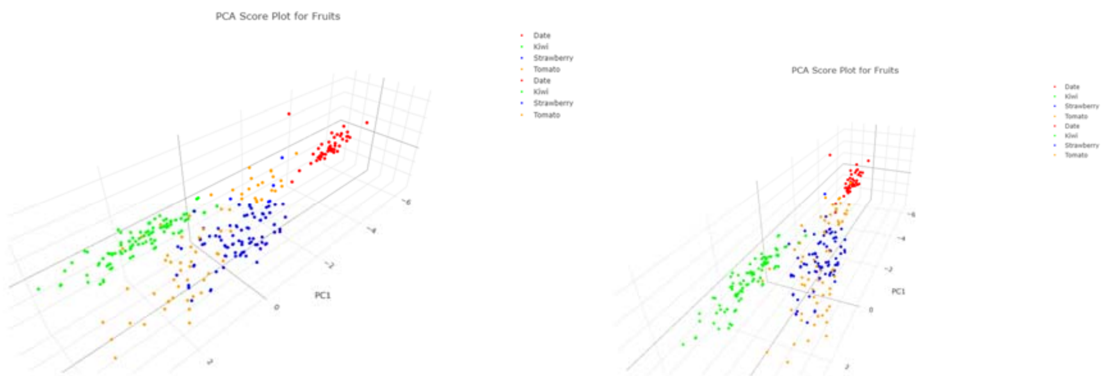


Figure 12 (b): PCA Analysis results

5.1.2 Classification Results

We developed two separate models for fruit classification and water-milk percentage prediction using machine learning algorithms. The dataset was collected manually and analyzed using the "mlr" and "ranger" packages in R. We used the makeClassifTask () function to convert the dataset into a classification task and the classif.ranger algorithm as the learning method. The mtry and num.trees hyperparameters were optimized using the tuneParams() function, and model performance was evaluated using cross-validation with accuracy (acc) and misclassification error (mmce) as evaluation metrics.

For fruit classification prediction, the tuned hyperparameters were found to be mtry = 1 and num.trees = 143, achieving an accuracy of 0.964. The cross-validation analysis showed an average accuracy of 0.955 and an average misclassification error of 0.045, suggesting that the classif.ranger algorithm can accurately classify fruits based on their spectral reflectance values.

Similarly, for water-milk percentage prediction, we used a cross-validation strategy with five iterations and evaluated the model's performance using accuracy and misclassification error measures. The tuned classif.ranger model achieved an accuracy of 99.09% on the training set and an accuracy of 96.36% on the test set, indicating good generalization performance. The mean misclassification error was 3.64%, suggesting that the model made relatively few errors in classifying the samples. These results demonstrate that the classif.ranger model with the optimal hyperparameters identified by the tuneParams function can accurately classify samples of milk based on their spectral reflectance data.

The results also suggest that machine learning algorithms, specifically the classif.ranger algorithm, can be effectively used for fruit classification and water-milk percentage prediction based on their spectral reflectance values. These models have the potential to be useful tools for quality control and food safety appl

5.2 Discussion:

The results of this study have significant implications for the agriculture and dairy industries. The use of NIR-based techniques and machine learning algorithms can help automate the classification process and provide a faster and more accurate way to grade fruits and milk. [6] The models developed for fruit classification and milk grading have shown promising results, with high accuracy rates. This could be useful for farmers to sort fruits based on their quality and ripeness, and for milk processors to ensure the appropriate quality and safety standards are met, reducing the risk of milk adulteration and improving public health.

However, it is important to note that the sample size used in this study was relatively small, and there is a need to test the models on larger datasets to determine their accuracy and reliability. Additionally, further research is required to test the models in real-world settings, where there may be variations in the spectral data due to environmental factors.

In terms of future research, there is a need to investigate the use of other machine learning algorithms, such as deep learning, for fruit classification and milk grading. Furthermore, the use of hyperspectral imaging could provide more detailed information about the composition of the samples, which could improve the accuracy of the models.

The results of this study have significant implications for the fields of fruit classification and milk grading. The use of NIR-based techniques, coupled with machine learning algorithms, can help automate the classification process and provide a faster and more accurate way to grade fruits and milk.

The findings suggest that the machine learning models developed in this study can accurately classify different types of fruits based on their spectral data. This could be useful for fruit sorting in the agricultural industry, enabling farmers to sort fruits based on their quality and ripeness, thereby improving the overall quality of the produce.

Similarly, the models developed for milk grading could be used to ensure that milk is of the appropriate quality and safety standards, thereby reducing the risk of milk adulteration and improving public health. This could be particularly important in developing countries, where milk adulteration is a growing health hazard. However, it is important to note that there are limitations to this study. The sample size used in this study was relatively small, and there is a need to test the models on larger datasets to determine their accuracy and reliability. Furthermore, the study was conducted under controlled laboratory conditions, and it is important to determine whether the models can perform well in real-world settings where there may be variations in the spectral data due to environmental factors. In terms of future research, there is a need to investigate the use of other machine learning algorithms, such as deep learning, for fruit classification and milk grading. [6] Additionally, there is a need to explore the use of hyperspectral imaging for these applications, as it has the potential to provide more detailed information about the composition of the samples.

Chapter 6: Conclusion

The NIR-based portable device was used to collect spectral data from fruit and milk samples, and for performing PCA analysis to visualize the clustering of the samples based on their properties. The PCA plots in Fig 12(a) and 12(b) showed that the collected data formed well-defined clusters, indicating that the proposed device has the potential to accurately distinguish between different fruits and detect different milk-water percentages. Specifically, for fruit classification, the authors found that the scores plot showed a good intra-cluster correlation for all the investigated fruit types, with kiwi showing the highest absorption levels across all wavelengths and dates showing the lowest absorption levels. For milk grading, the mean absorption levels decreased as the percentage of water added to the milk increased, with the PCA analysis showing distinct clusters formed by different milk-water solutions. These findings demonstrate the effectiveness of the proposed device for both fruit classification and milk-water percentage detection.

We have developed a low-cost portable device for fruit classification using NIR sensors AS7263 to measure the absorption of light by fruit samples at six different wavelengths. The data obtained from these sensors are processed using machine learning algorithms to classify fruits into different categories based on their properties. Our device is designed to be accessible and affordable for researchers who cannot afford expensive sorting machines, providing a cost-effective solution for fruit classification. The device also features a web app that allows real-time viewing of classification results, aiding informed decision-making. The device consists of hardware which is responsible for collecting reading and Web Application.

Separate machine learning models developed in backend for fruit classification and water-milk percentage prediction, using R showed that the classif.ranger algorithm can accurately classify fruits based on their spectral reflectance values, achieving an accuracy of 0.964, and could accurately classify samples of milk based on their spectral reflectance data, achieving an accuracy of 96.36%. These results suggest that machine learning algorithms can be effectively used for quality control and food safety applications in the agriculture and dairy industries.

These models when used in R Shiny app form autonomous and state of art Fruit Classification and Milk Adulteration grading systems. Use of the cloud has enabled remote and instant access to real time sensor reading. Data Collection features can be used by researchers to innovate new solutions using similar techniques. Moreover, the cost of development of such a device is comparatively very low as compared to already available market Alternatives and gives good performance. This work finds multiple applications in future to replace the costly NIR devices. According to Bernard G. Barthès, in his research paper "Performance comparison between a miniaturized and a conventional near infrared reflectance (NIR) spectrometer for characterizing soil carbon and nitrogen," the availability of spectrometers is often limited due to their high price. This limits the accessibility of NIR spectrometry to researchers and farmers, who could benefit from using this technology to assess soil quality. The high cost of spectrometers is primarily due to the complexity and precision of their components, which are necessary to accurately measure the reflectance spectra of samples. As a result, many researchers and farmers cannot afford to purchase a spectrometer, which can cost tens of thousands of dollars.

In this study the complex equipment was replaced with cloud-based processing. This not only made use of the device more economical but also robust to multiple applications. This research deals with applying two such applications. However, the existing infrastructure can be innovative further to inherit existing NIR based classification methods, providing a user-friendly approach to application development.

In conclusion, the use of a portable NIR-based device for fruit classification and milk grading has been shown to be effective in accurately distinguishing between different fruits and detecting different milk-water percentages. Machine learning algorithms have also been demonstrated to be effective for quality control and food safety applications in the agriculture and dairy industries. The development of a cost-effective, miniaturized NIR-based device, as demonstrated in this study, has the potential to replace the costly spectrometers currently available in the market, which can be limiting for researchers and farmers. By leveraging cloud-based processing, this device not only provides an economical alternative to expensive spectrometers, but also enables researchers and farmers to access real-time sensor data remotely. Overall, this study presents a promising approach to the development of low-cost, portable NIR-based devices for various applications in the food and agriculture industries.

.

Chapter 7: Future Work

There are several avenues for future research that could build on the findings of this study and further improve the efficiency and accuracy of the proposed device for food quality assessment. In this section, we outline some potential areas for future work:

7.1 Testing the device in real-world settings

One limitation of this study is that it was conducted under controlled conditions with constraints. Further research is needed to determine whether the proposed device can perform well in real-world settings, where there may be variations in the spectral data due to environmental factors. Field trials could be conducted to assess the device's accuracy in different environmental conditions and to determine whether it can be effectively used by farmers, food processors, and retailers.

7.2 Developing models for other food items

In this study, we focused on fruit classification and milk grading. However, the proposed device could be used for the quality assessment of other food items, such as vegetables, grains, and meat products. Future research could explore the feasibility of using the device and machine learning algorithms for these applications.

7.3 Integration with blockchain technology

Blockchain technology has the potential to improve the transparency and traceability of food products. Future research could explore the integration of the proposed device with blockchain technology to create a more transparent and trustworthy food supply chain. This could include the development of a blockchain-based system for storing and sharing food quality data, which could be accessed by all stakeholders in the supply chain.

7.4 Collaborations with food industry stakeholders

Collaboration with food industry stakeholders, such as farmers, food processors, and retailers, could help to ensure that the proposed device meets their needs and is tailored to their specific requirements. Future research could involve collaboration with these stakeholders to identify key performance indicators and to develop customized applications for the device.

7.5 Application in other fields

The proposed device and machine learning algorithms could be applied to other fields beyond the food industry, such as environmental monitoring and medical diagnostics. Future research could explore the feasibility of using the device for these applications and develop customized models for different fields.

- **Environmental monitoring:** The device could be used to analyze the composition of soil, water, and air samples, allowing for more accurate and efficient monitoring of environmental pollutants and contaminants. Machine learning models could be developed to classify different types of pollutants and predict their concentrations based on spectral data.

- **Medical diagnostics:** The device could be used to analyze biological samples, such as blood or urine, to diagnose diseases and monitor treatment efficacy. Machine learning models could be developed to identify biomarkers of various diseases and predict patient outcomes based on spectral data.
- **Forensic analysis:** The device could be used to analyze physical evidence, such as fibers or fingerprints, to identify suspects and solve crimes. Machine learning models could be developed to classify different types of evidence and match them to potential suspects based on spectral data.
- **Quality control in manufacturing:** The device could be used to analyze the composition of raw materials and finished products in manufacturing industries, such as pharmaceuticals or electronics. Machine learning models could be developed to detect defects or contaminants and predict product performance based on spectral data.
- **Agriculture and forestry:** The device could be used to analyze the composition of soil and plant samples, allowing for more efficient and accurate monitoring of crop health and nutrient levels. Machine learning models could be developed to predict crop yields and identify potential problems based on spectral data. Additionally, the device could be used to analyze the composition of wood samples for forest management and product development purposes.

References and Work Cited

1. Furhan Iqbal. (2019). Milk Adulteration: A Growing Health Hazard in Pakistan. In H. Siddiqi & K. Tariq (Eds.), Emerging Trends in Dairy Science and Dairy Products (pp. 319-329). Springer. https://doi.org/10.1007/978-3-030-18668-4_17

2. Zeb, A., Qureshi, W. S., Ghafoor, A., & O'Sullivan, D. (2019). Learning fruit class from short wave near infrared spectral features, an AI approach towards determining fruit type. Journal of King Saud University-Computer and Information Sciences, 31(4), 519-525. DOI: 10.1016/j.jksuci.2018.06.00118.

3. B. G. Barthès, "Performance comparison between a miniaturized and a conventional near-infrared reflectance (NIR) spectrometer for characterizing soil carbon and nitrogen," Geoderma, vol. 326, pp. 43–50, 20

4. Manley M, Baeten V. Spectroscopic technique: Near infrared (NIR) spectroscopy. InModern techniques for food authentication 2018 Jan 1 (pp. 51-102). Academic Press.

5. Wu, W., Massart, D. L., & De Jong, S. (1997). Kernel-PCA algorithms for wide data Part II: Fast cross-validation and application in classification of NIR data. Chemometrics and Intelligent Laboratory Systems, 37(2), 271-280.

6. Galata, D. L., Könyves, Z., Nagy, B., Novák, M., Mészáros, L. A., Szabó, E., ... & Nagy, Z. K. (2021). Real-time release testing of dissolution based on surrogate models developed by machine learning algorithms using NIR spectra, compression force and particle size distribution as input data. International Journal of Pharmaceutics, 597, 120338.

7. Beć, K. B., Grabska, J., & Huck, C. W. (2021). Principles and applications of miniaturized near-infrared (NIR) spectrometers. Chemistry–A European Journal, 27(5), 1514-1532.

8. Kondo, N. (2010). Automation on fruit and vegetable grading system and food traceability. Trends in Food Science & Technology, 21(3), 145-152.

Advance Food Grading FYP Thesis' rakhna

ORIGINALITY REPORT

6%	4%	3%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet Source	1%
2	Submitted to Higher Education Commission Pakistan Student Paper	1%
3	etd.aau.edu.et Internet Source	<1%
4	www.sparkfun.com Internet Source	<1%
5	M. Anandarajan. "Glaucoma screening: a web-based intelligent system approach", International Journal of Healthcare Technology and Management, 2001 Publication	<1%
6	Ximena Omlin, Francesco Crivelli, Monika Näf, Lorenz Heinicke et al. "The Effect of a Slowly Rocking Bed on Sleep", Scientific Reports, 2018 Publication	<1%