A Project Report

on

Milk Quality Monitoring System Using IoT And ML

Submitted in partial fulfillment of the requirements

for the award of degree of

BACHELOR OF TECHNOLOGY

in

Information Technology

by

Gaddamedi Harshitha (20WH1A1271)

Kandela Kavitha (20WH1A1276)

Nampally Manjula(20WH1A1286)

Under the esteemed guidance of

Ms. Ch. Sai Lalitha Bala

Assistant Professor

Department of Information Technology

BVRIT HYDERABAD College of Engineering for Women

Rajiv Gandhi Nagar, Nizampet Road, Bachupally, Hyderabad – 500090

(Affiliated to Jawaharlal Nehru Technological University, Hyderabad)

(NAAC ‘A’ Grade & NBA Accredited- ECE, EEE, CSE & IT)

June, 2024

DECLARATION

We hereby declare that the work presented in this project entitled **"Milk Quality Monitoring System Using IoT And ML"** submitted towards completion of IV year II sem of B.Tech IT at "BVRIT HYDERABAD College of Engineering for Women", Hyderabad is an authentic record of our original work carried out under the esteemed guidance of Ms. Ch. Sai Lalitha Bala, Assistant Professor, Department of Information Technology.

Gaddamedi Harshitha (20WH1A1271)

Kandela Kavitha (20WH1A1276)

Nampally Manjula(20WH1A1286)

BVRIT HYDERABAD

College of Engineering for Women

Rajiv Gandhi Nagar, Nizampet Road, Bachupally, Hyderabad – 500090

(Affiliated to Jawaharlal Nehru Technological University Hyderabad)

(NAAC ‘A’ Grade & NBA Accredited- ECE, EEE, CSE & IT)

CERTIFICATE

This is to certify that the Project report on **“ Milk Quality Monitoring System Using IoT And ML”** is a bonafide work carried out by **G. Harshitha (20WH1A1271), K. Kavitha(20WH1A1276), N. Manjula(20WH1A1286)** in the partial fulfillment for the award of B.Tech degree in **Information Technology, BVRIT HYDERABAD College of Engineering for Women, Bachupally, Hyderabad** affiliated to Jawaharlal Nehru Technological University, Hyderabad under my guidance and supervision.

The results embodied in the project work have not been submitted to any other university or institute for the award of any degree or diploma.

Internal Guide
Ms. Ch. Sai Lalitha Bala
Assistant Professor
Department of IT

Head of the Department
Dr. Aruna Rao S L
Professor & HoD
Department of IT

External Examiner

ACKNOWLEDGEMENT

We would like to express our profound gratitude and thanks to **Dr. K. V. N. Sunitha, Principal, BVRIT HYDERABAD College of Engineering for Women** for providing the working facilities in the college.

Our sincere thanks and gratitude to **Dr. Aruna Rao S L, Professor & Head, Department of IT, BVRIT HYDERABAD College of Engineering for Women** for all the timely support, constant guidance and valuable suggestions during the period of our project.

We are extremely thankful and indebted to our internal guide, **Ms. Ch. Sai Lalitha Bala, Assistant Professor, Department of IT, BVRIT HYDERABAD College of Engineering for Women** for her constant guidance, encouragement and moral support throughout the project.

Finally, we would also like to thank our Project Coordinators **Dr. J. Kavitha, Associate Professor and Mr. Mukhtar Ahmad Sofi, Assistant Professor**, all the faculty and staff of Department of IT who helped us directly or indirectly, parents and friends for their cooperation in completing the project work.

Gaddamedi Harshitha (20WH1A1271)

Kandela Kavitha (20WH1A1276)

Nampally Manjula (20WH1A1286)

ABSTRACT

The dairy industry relies on maintaining consistent milk product quality to meet consumer demands and ensure food safety standards. To address this concern, a Milk Quality Monitoring System (MQMS) integrates ”The convergence of the Internet of Things (IoT) with Machine Learning technologies”. By deploying IoT sensors throughout dairy farms and processing facilities, the MQMS continuously gathers data on critical milk quality parameters such as temperature, pH, fat content, and color. These data streams are transmitted in real time to a central server for analysis. Unlike conventional methods that rely solely on taste, temperature, and turbidity, which often yield inaccurate results, incorporating additional parameters like fat content and color enhances milk quality testing efficiency. The MQMS guarantees consistent milk quality, mitigating the risk of substandard products reaching consumers and thereby improving food safety standards and consumer confidence through accurate classification of milk quality. Furthermore, the system’s capability to detect contamination and adulteration leads to cost savings for dairy owners.

Keywords: Dairy industry, Milk Quality Monitoring System (MQMS)’, Sensors, Temperature, pH, Fat content, Color, Contamination, Cost savings.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1. INTRODUCTION

The integration of IoT (Internet of Things) technology into milk quality monitoring systems is a significant leap forward for the dairy industry. IoT sensors are strategically placed throughout the milk production and distribution network to gather real-time data on critical parameters such as temperature, humidity, and acidity. These sensors ensure continuous monitoring, providing up-to-the-minute information on the conditions that affect milk quality. By detecting any deviations from the optimal conditions, IoT technology helps prevent the spoilage of milk, ensuring that only high-quality products reach the consumer.

The use of IoT in milk quality monitoring extends beyond simple data collection. IoT devices can be embedded in milking machines and storage units, providing detailed insights into every stage of milk production. For example, temperature sensors in milk tanks ensure that the milk is stored at the correct temperature, while pH sensors monitor acidity levels to detect any signs of contamination. Additionally, the integration of RFID (Radio-Frequency Identification) tags ensures robust traceability, allowing producers to track each batch of milk from the farm to the consumer. This level of traceability not only enhances transparency across the supply chain but also helps in quickly identifying and addressing any quality issues that may arise.

Moreover, IoT technology can facilitate predictive maintenance of equipment used in dairy production. By continuously monitoring the operational status of milking machines, refrigeration units, and other critical machinery, IoT sensors can detect early signs of wear and tear or malfunction. This enables dairy farmers to perform maintenance before a breakdown occurs, reducing downtime and preventing potential losses due to equipment failure.

The integration of IoT also supports environmental sustainability within the dairy industry. Sensors can monitor water usage, energy consumption, and waste production, providing data that can be used to optimize resource management. For instance, IoT-enabled systems can automate the adjustment of cooling systems in response to real-time temperature changes, reducing energy consumption and minimizing the carbon footprint of dairy operations.

Additionally, IoT devices can enhance animal welfare by monitoring the health and behavior of dairy cows. Wearable sensors can track vital signs, movement patterns, and feeding behavior, providing early warnings of health issues such as infections or lameness. This real-time health monitoring allows for timely veterinary interventions, improving animal health and productivity.

In the context of regulatory compliance, IoT technology ensures that dairy operations adhere to stringent standards for food safety and quality. Automated data logging and reporting streamline the documentation process, making it easier for producers to meet regulatory requirements and pass inspections. This not only protects consumers but also enhances the reputation and marketability of dairy products. Machine Learning (ML) algorithms play a crucial role in the Milk Quality Monitoring System by analyzing the vast amounts of data collected by IoT sensors. These algorithms are designed to recognize patterns, identify anomalies, and predict potential quality issues. By scrutinizing data on temperature, humidity, acidity, and other variables, ML algorithms can detect subtle changes that may indicate a problem, such as an increase in bacterial content or deviations from the optimal storage conditions.

The predictive capabilities of ML algorithms enable the system to anticipate issues before they become critical, allowing for rapid interventions. For instance, if the system detects a trend that suggests a possible future contamination, it can alert dairy farmers and producers, enabling them to take preventive measures. This proactive approach not only helps maintain high-quality milk production but also reduces waste and economic losses by preventing the distribution of inferior or contaminated milk.

Moreover, ML algorithms can enhance the efficiency of dairy operations by minimizing manual monitoring tasks. By automating the data analysis process, these algorithms free up time for dairy farmers and producers, allowing them to focus on other critical aspects of their operations. The scalability of ML-driven monitoring processes means that even large-scale dairy operations can maintain consistent quality standards without a proportional increase in manual labor.

In addition to operational efficiency, ML algorithms contribute to better decision-making processes. By providing detailed insights and actionable recommendations based on data analysis, ML models can guide dairy farmers in optimizing feeding schedules, managing herd health, and improving overall farm management. This data-driven approach allows for more informed decisions that can enhance productivity and profitability.

ML algorithms also facilitate continuous improvement through their ability to learn and adapt over time. As more data is collected, these algorithms refine their models, improving the accuracy of predictions and anomaly detection. This continuous learning cycle ensures that the Milk Quality Monitoring System remains effective and up-to-date with the latest trends and patterns in milk production.

Furthermore, the integration of ML with IoT in milk quality monitoring supports compliance with food safety regulations. Automated detection of quality deviations and real-time reporting ensure that dairy products meet stringent safety standards, reducing the risk of regulatory breaches and associated penalties. This compliance not only protects consumers but also strengthens the trust and credibility of dairy brands in the market.

Another significant benefit of ML algorithms is their role in enhancing supply chain transparency. By analyzing data from various stages of production and distribution, ML models can provide a comprehensive overview of the entire supply chain. This visibility helps identify bottlenecks, inefficiencies, and potential risks, enabling stakeholders to optimize logistics and ensure timely delivery of high-quality milk products.

1.1. Motivation

The motivation for predicting milk quality using IoT (Internet of Things) and ML (Machine Learning) lies in the pursuit of enhancing dairy industry efficiency, product quality, and overall sustainability. By leveraging IoT devices such as sensors and smart devices installed in milking machines and storage units, real-time data on various parameters such as temperature, pH levels, and bacterial content can be collected. Machine Learning algorithms can then analyze this data to identify patterns, detect anomalies, and predict potential issues in milk quality. This proactive approach allows dairy farmers and producers to take immediate corrective actions, ensuring the production of high-quality milk while minimizing waste and economic losses. Additionally, the integration of IoT and ML enables automated monitoring, reducing the manual workload and increasing the scalability of dairy operations. Ultimately, this convergence of technology aims to optimize milk production processes, guarantee product safety, and contribute to a more sustainable and technologically advanced dairy industry.

1.2. Objective

The primary objective of implementing IoT and ML for the prediction of milk quality is to revolutionize the dairy industry by enhancing operational efficiency, ensuring product quality, and promoting sustainability. By deploying IoT devices, such as sensors, across the milk production and storage chain, real-time data on crucial parameters like temperature, pH, and bacterial content can be collected. The integration of Machine Learning algorithms then allows for comprehensive analysis, pattern recognition, and early detection of potential quality issues. This proactive approach empowers dairy farmers to take immediate corrective actions, minimizing wastage and economic losses, while simultaneously optimizing the production process. Furthermore, the automated monitoring capabilities provided by IoT and ML not only streamline operations but also enable scalability in dairy farming. The motivation for employing these technologies is rooted in the desire to ensure the production of high-quality milk, aligning with consumer expectations and regulatory standards. Additionally, the reduction of manual workload through automation contributes to the overall sustainability of dairy operations. By embracing IoT and ML, the dairy industry can advance towards a more intelligent and technologically sophisticated future, promoting resource efficiency and environmental responsibility. Ultimately, the objective is to create a resilient, high-quality dairy ecosystem that not only meets the demands of the modern market but also addresses global challenges by minimizing the ecological footprint associated with dairy production.

1.3. Problem Definition

The challenge in predicting milk quality through ML and IoT lies in the dairy industry's need for real-time, accurate assessments. The traditional manual methods for evaluating milk quality are time-consuming and susceptible to errors, hindering timely interventions and decision-making. The absence of continuous monitoring exacerbates the difficulty in identifying deviations in parameters like fat content and bacterial load, which are crucial for ensuring product quality. Integrating ML algorithms with IoT sensors is necessary to establish a predictive system capable of monitoring and analyzing key quality indicators in real-time. This involves overcoming the limitations of current practices, enhancing precision, and providing dairy farmers with actionable insights for immediate responses to quality fluctuations. The goal is to develop an innovative solution that transforms the dairy industry by ensuring consistent, high-quality milk production, reducing waste, and promoting sustainability through data-driven decision support.

1.4. Proposed Method

The Milk Quality Prediction System represents a groundbreaking approach in ensuring the highest standards of milk quality throughout its production and distribution processes. By harnessing the power of IoT sensors, microcontrollers like Arduino Uno, and an array of specialized components such as temperature, color, and pH sensors, the system can meticulously monitor various parameters critical to milk quality in real-time. This constant surveillance allows for the early detection of any deviations from the desired standards, enabling swift corrective actions to be taken.

Moreover, the integration of machine learning algorithms enhances the system's capabilities by enabling it to analyze vast amounts of data collected from the sensors. These algorithms can identify patterns and correlations that might indicate potential quality issues before they manifest, providing invaluable insights for preemptive measures. Additionally, the predictive analytics component of the system enables it to forecast potential quality fluctuations based on historical data and current trends, further optimizing proactive interventions.

1.5. Dataset

The data source is the file milk.csv. It contains the data for this example in comma-separated values (CSV) format. The number of columns is 8, and the number of rows is 1059.The dataset split into training and test sets as follows:

Training set: 741 rows

Test set: 318 rows

The variables are:

pH: This feature defines pH of the milk, which is in the range of 3 to 9.5.

temperature: This feature defines the temperature of the milk, and its range is from 34'C to 90'C.

taste: This feature defines the taste of the milk and takes the possible values: 1 (good) or 0 (bad).

odor: This feature defines the odor of the milk and takes the possible values: 1 (good) or 0 (bad).

fat: This feature defines the fat of the milk and takes the possible values: 1 (good) or 0 (bad).

turbidity: This feature defines the turbidity of the milk and takes the possible values: 1 (good) or 0 (bad).

colour: This feature defines the color of the milk, which is in the range of 240 to 255.

grade: This is the target and takes the values low quality, medium quality, or high quality.

The milk dataset contains 429 instances of low quality, 374 instances of medium quality, and 256 instances of high quality.

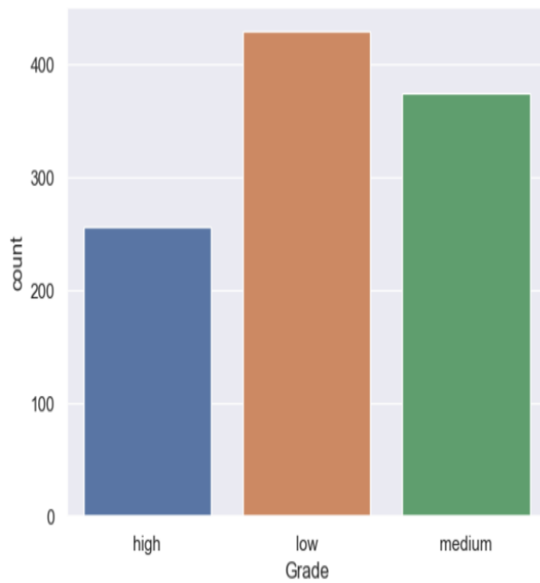

Figure 1.5.1: Classification of data

2. IoT COMPONENTS

2.1. pH Sensor:

The pH sensor's integration with IoT and ML revolutionizes milk quality prediction. Continuously measuring acidity, the sensor transmits real-time data, enabling ML algorithms to identify deviations and predict spoilage or contamination. The proactive approach ensures high-quality milk production and distribution, with real-time monitoring enhancing efficiency and waste reduction benefits for stakeholders in the dairy industry. The pH of fresh, unprocessed milk typically ranges between 6.5 and 6.7. Milk spoilage occurs when lactic acid production by bacteria decreases its pH, indicating spoilage and prompting intervention if the pH drops below 6.5.

Figure 2.1.1: pH Sensor

2.2. DS18B20 Temperature Sensor:

The DS18B20 temperature sensor is integral to milk quality monitoring systems, precisely measuring temperature in dairy storage. This sensor ensures optimal conditions, preventing bacterial growth and maintaining freshness. It enhances quality control by providing real-time data on temperature fluctuations. Specifically designed for accuracy and reliability, the DS18B20 enables dairy producers to adhere to stringent quality standards, ensuring consumers receive safe, high-quality milk products while optimizing production processes and minimizing waste in the dairy industry. Raw milk storage should be kept between 1°C to 4°C to prevent bacterial growth and spoilage. Temperatures below 1°C can cause freezing, affecting milk texture and quality. Regulatory standards, like the FDA, recommend storing milk at 45°F or below to maintain quality and safety.

Figure 2.2.1: Temperature Sensor

2.3. Lactometer:

The lactometer, when combined with machine learning and IoT, can predict milk quality by assessing milk composition, detecting potential adulteration, and identifying quality deviations, thereby enhancing efficiency and consumer trust in the dairy industry. Lactometers measure milk density, fat content, and protein content, detect adulteration, set thresholds, and set quality indicators. Machine learning algorithms analyze data, validate readings, and real-time monitoring via IoT allows immediate action.

Figure 2.3.1: Lactometer

2.4. TCS230 Color Sensor :

The TCS230 color sensor is pivotal in milk quality monitoring systems, detecting variations in milk color that may indicate spoilage or irregularities. Analyzing color changes helps identify potential quality issues, ensuring consumers receive fresh and safe dairy products. The TCS230 enhances quality control measures by providing rapid and accurate color readings, enabling dairy producers to take corrective actions promptly. This sensor contributes to maintaining high standards in milk production, promoting consumer satisfaction, and minimizing risks associated with sub-optimal product quality. Establish baseline color profiles for high-quality milk, monitor changes for spoilage, and use machine learning analysis to identify patterns, set thresholds, trigger alerts, and validate sensor readings. This feature defines the color of the milk, which is in the range of 240 to 255.

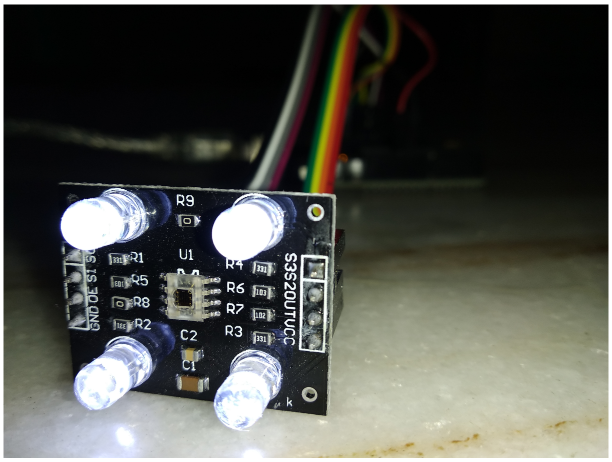

Figure 2.4.1: Color Sensor

3. LITERATURE SURVEY

IoT-based detection of adulteration in milk:

[1] It revolves around the development of an Internet of Things (IoT)-based system designed for detecting adulteration in milk. This innovative system employs sensors to measure diverse parameters, including air quality, conductivity, pH, and temperature, to comprehensively assess the quality of milk. The execution of this proposed framework is facilitated by the Arduino UNO microcontroller. The sensor readings are then showcased on an LCD screen, and an Android app named Blynk is utilized for real-time monitoring. To enhance communication, the system incorporates a GSM modem interface, enabling the transmission of sensor values via SMS to a predefined mobile number. The study underscores the significance of ensuring the purity of milk and highlights potential health risks associated with the consumption of adulterated milk. Tests were conducted using both pure and adulterated milk samples, with adulterants such as detergent. The results were prominently displayed on both the LCD screen and the Blynk application, showcasing the system's efficacy in detecting adulteration. The research concludes by emphasizing the system's potential applications in both small and large-scale milk dairies. Additionally, it explores future prospects for integrating IoT and database management systems to enhance milk quality assessment and streamline payment processing. Furthermore, the study references pertinent literature on milk spoilage detection, particularly using remote-query sensors.IoT-based system emerges as a promising solution for ensuring the quality of milk, with real-time monitoring capabilities and the potential for widespread implementation across various dairy settings. The integration of advanced technologies not only addresses the issue of adulteration but also opens avenues for further research and innovation in the domain of milk quality assurance and management.

Research on dairy products based on machine learning:

[2] Approach utilizing an electronic nose model and machine learning algorithms for dairy product detection is presented. The electronic nose, mimicking the olfactory organs of animals, is equipped with a gas sensor array, signal acquisition module, data acquisition module, and signal processing and pattern recognition module. The experiment involved collecting volatile gases from milk samples using the electronic nose, with subsequent data analysis conducted on 1000 standardized data groups, comprising 800 training data and 200 test data. To reduce data dimensionality, principal component analysis (PCA) and linear discrimination analysis (LDA) were employed for cattle farm classification. The machine learning algorithms Logistic Regression, Support Vector Machine, and Random Forest were utilized for classification and regression tasks, demonstrating their effectiveness in distinguishing milk sources, estimating milk fat and protein content, and evaluating milk quality. The study's findings highlight the potential of machine learning algorithms in dairy product detection, offering a low-cost, non-destructive, and accurate method for assessing milk quality and authenticity. This research contributes to the advancement of technology in the dairy industry, with implications for enhancing quality control and ensuring consumer satisfaction.

Method	**Accuracy**
Logistic Regression	91.5 %
Support Vector Machine	96 %
Random Forest	96 %

3.2.1: Performance Analysis of used methods

Machine Learning Applied to Milk Sample Classification:

[3] The application of machine learning techniques to classify milk samples, with the goal of ensuring the quality and safety of dairy products for human consumption. Emphasizing the significance of providing high-quality dairy products, particularly for vulnerable populations like children, the research evaluates the classification process using machine learning models such as random forest, k-nearest neighbors (KNN), and neural networks. The results reveal that the random forest technique achieved the highest accuracy, ranging from 96 % to 99 %, while KNN attained an accuracy level between 77 % and 81 %. The study also explores the use of support vector machines and linear discriminant analysis for classification, as well as the potential application of e-nose and brix meters for food quality monitoring. The findings of this research contribute to the advancement of techniques for ensuring the integrity and safety of dairy products, with potential implications for the food industry and public health. The study's focus on leveraging machine learning to enhance the classification of milk samples underscores the importance of technological advancements in ensuring the quality and safety of food products, aligning with global efforts to improve food security and public health.

Method	**Accuracy**
Support Vector Machine	90 %
Random Forest	98 %

3.3.1: Analysis of used methods

Using Machine Learning Algorithms to Detect Milk Quality:

[4] Üsing Machine Learning Algorithms to Detect Milk Quality"presents a study on the application of machine learning algorithms to determine milk quality. The research utilizes the Milk Quality dataset from Kaggle, comprising seven attributes including pH, temperature, taste, odor, fat, turbidity, and color. By employing the AdaBoost and Neural Network algorithms, the study aims to classify milk quality as low, medium, or high. The results demonstrate high success rates, with the AdaBoost algorithm achieving a 99.9 % success rate and the Neural Network algorithm achieving 95.4 %. The study emphasizes the potential of machine learning in enhancing food safety and quality control in the dairy industry. It highlights the importance of analyzing multiple features to accurately assess milk quality, offering promising implications for improving milk quality assessment. The research has the potential to revolutionize milk quality monitoring, benefiting both food producers and consumers. By leveraging intelligent systems, this approach could contribute to the advancement of dairy product quality and have a significant impact on the milk production and consumption landscape. Overall, the study underscores the potential of machine learning algorithms in ensuring the quality and safety of dairy products, with implications for the broader food industry.

Method	Accuracy
AdaBoost	99.9%
Neural Network	95.4%

3.4.1: Analysis of used methods

An analysis and design of fresh milk smart grading system based on the Internet of Things:
[5] It presents an analysis and design of a fresh milk smart grading system based on the Internet of Things (IoT). The study aims to improve the quality classification in the dairy industry through the use of machine learning models. The system involves monitoring the temperature of fresh milk in real-time using IoT sensors and collecting quality data to be classified by an artificial neural network (ANN) model. The study also uses Business Process Modelling Notation (BPMN) diagrams to describe the process and sub-process in a smart grading system of fresh milk based on IoT. The result of the requirement analysis showed how smart the grading system involved stakeholders. The machine learning model can help the IoT system classify goods or services. The accuracy value of the classification obtained is 98.74%. The attributes used for grouping were temperature and color. The study provides insights into the potential of IoT and machine learning in improving the quality classification of fresh milk in the dairy industry.

Smart system for Milk quality analysis and billing system:
[6] A smart system for milk quality analysis and billing system, which aims to improve the dairy sector in India's economy. The system measures various parameters of milk, including fat content, harmful gas content, pH level, and refrigerated level, using sensors such as LDR and pH sensors. The data is then transmitted to an IoT cloud platform and displayed on LCD monitors, allowing for internet-based data monitoring. The system also includes a Blynk app on an Android phone to aid in billing computations and daily payments. The paper highlights the importance of technology in improving the agricultural way of life and offers a solution that can benefit small-scale farmers in India. Overall, the smart system for milk quality analysis and billing system offers a comprehensive

solution to improve the dairy sector's efficiency and productivity.

Predicting cow milk quality traits from routinely available milk spectra using statistical machine learning methods:

[7] It mainly discusses an IoT-based milk monitoring system designed to ensure the quality and quantity of milk, particularly in rural and urban areas. The system aims to address the issue of milk adulteration and the presence of pathogenic organisms in raw milk, which can lead to health risks and a decline in the quality of life. The proposed system utilizes various sensors such as milk level, gas, temperature, viscosity, and salinity sensors, along with an Arduino controller for real-time monitoring and quality checking. Additionally, the system includes power analysis to assess power consumption and battery backup requirements. The implementation results indicate that the system can effectively detect milk adulteration and early microbial activity, contributing to the overall safety and quality of milk products. The use of IoT technology allows for remote monitoring and integration of the physical environment into computer-based systems, leading to improved efficiency, accuracy, and economic benefits. However, challenges related to sensor power consumption and system optimization are also addressed. Overall, the IoT-based milk monitoring system presents a promising approach to ensuring the safety and quality of milk products, with potential benefits for both producers and consumers.

Method	**Accuracy**
Support Vector Machine(SVM)	75%

3.7.1: Performance of listed methods

IoT for Development of Smart Dairy Farming:

[8] It explores the potential of the Internet of Things (IoT) and data-driven techniques in smart dairy farming (SDF) to improve milk yield and meet the increasing demand for dairy products. The authors propose a framework for enhancing milk production through the adoption of the latest techniques for improving feeding and milking procedures. The framework involves the use of wearable sensors to capture data from cows, which is then transferred to a base station and analyzed in the cloud using IoT-based platforms. The research emphasizes the potential of IoT-enabled smart dairy farming to address traditional farming challenges and increase milk production. Additionally, the article discusses the benefits of using AI techniques, such as smart monitoring, cow observation, feeding, milking, and reproduction management, in dairy farming. The proposed system is expected to make IoT-based farming more efficient, although it may require heavy initial investment. Overall, the research aims to assist farmers in increasing milk production while balancing the investment and earnings. The findings of the research have implications for both developed and developing countries in the dairy industry, offering potential solutions to improve production and address challenges in dairy farming.

IoT-based milk monitoring system for detection of milk adulteration:

[9] It presents an IoT-based milk monitoring system that utilizes various sensors to detect adulteration and ensure the quality and safety of milk. The system includes a Milk Level Sensor, Gas Sensor, Temperature Sensor, Viscosity Sensor, and Salinity Sensor, which continuously monitors the milk's composition and detects any early microbial activity or adulteration. The Arduino controller processes the data from these sensors and displays the results on an LCD screen. The system's power consumption is analyzed, and the authors emphasize the importance of real-time monitoring and decision-making to ensure the quality and safety of milk. The proposed system offers a promising solution for addressing milk quality control and safety concerns, while also highlighting the significance of IoT in enhancing efficiency, accuracy, and financial benefits in milk monitoring.

Near-infrared spectroscopic sensing system for online milk quality assessment in a milking robot:

[10] The development and validation of a near-infrared spectroscopic sensing system for online milk quality assessment in a milking robot. The system allows for the automatic collection of NIR spectra of raw milk, enabling the determination of major milk constituents, somatic cell count, and milk urea nitrogen. The precision and accuracy of the calibration models have been validated, with impressive results. The system has the potential to revolutionize milk quality assessment in the dairy industry, improving production efficiency and reducing costs. It provides detailed information on the system's specifications, including the spectrum sensor, light source, optical fiber, milk chamber surface, spectrometer, and data processing computer. The potential implications of this technology for the dairy industry are significant, as it could lead to more efficient and accurate milk quality assessment, ultimately benefiting both producers and consumers.

4. SYSTEM ARCHITECTURE

ML-IoT architecture in dairy farms deploys sensor networks for real-time data collection on quality parameters like fat content and bacterial load. Data is transmitted to a centralized cloud platform where ML algorithms analyze and forecast milk quality. Results are accessible via user-friendly interfaces, aiding timely decision-making. This integration optimizes quality control, reduces wastage, and enhances operational efficiency, fostering sustainable production of high-quality dairy products.

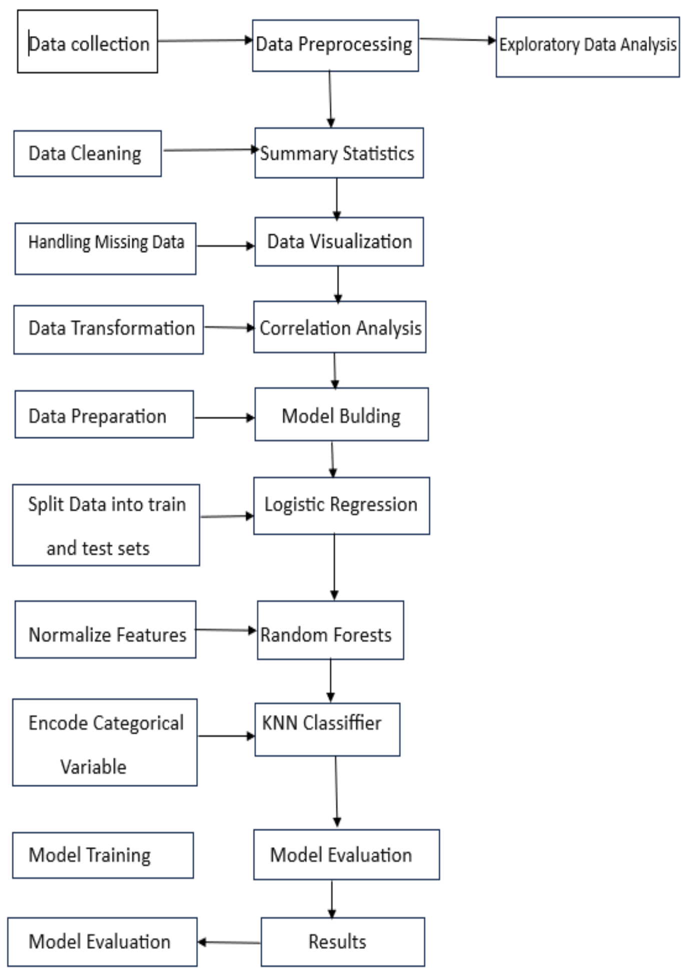

Figure 4.1: Architecture

5. IMPLEMENTATION

The pH sensor has the following connections: the output pin (o/p) should be connected to the A0 (analog pin of Arduino), the ground output pin (Gnd) to Gnd, the +9V pin to a separate 9V battery, and the ground pin to Gnd. The color sensor has the following connections: S0 to digital pin 8 of Arduino, S1 to digital pin 9, S2 to digital pin 12, S3 to digital pin 11, Sout to digital pin 10, OE to Gnd, Gnd to Gnd, and Vcc to +5V of the Arduino. The temperature sensor has the following connections: the red wire to +5V of the Arduino, the black wire to Gnd, and the yellow/white wire to A2 (analog pin of Arduino).

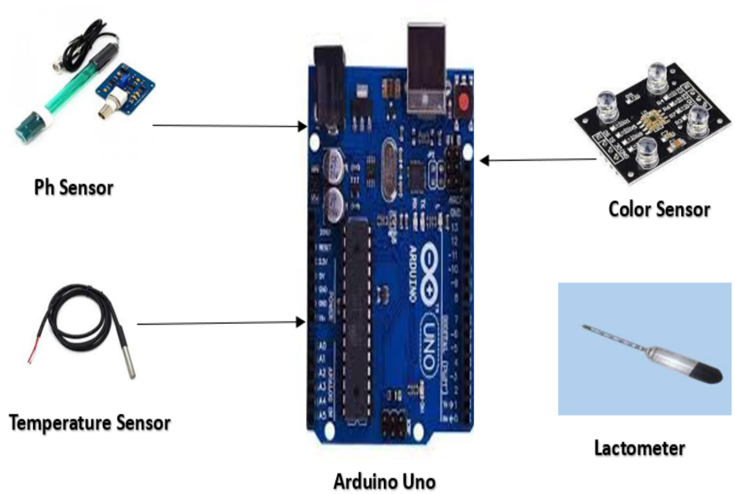

Figure 5.1: System design

5.1. Connections:

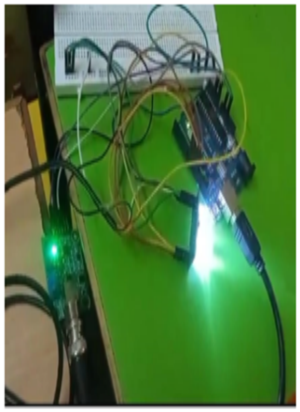

Figure: 1

Figure: 2

Figure 5.1.1: Connections

Figure 1: illustrates the connections between the Arduino Uno and the Ph Sensor.

Figure 2: The quality of milk is checked using a Ph Sensor.

6. EXISTING METHOD

Refined Feature Selection: Investigate additional relevant features that may impact milk quality and incorporate them into the predictive model. Conduct thorough feature engineering to ensure the inclusion of the most informative variables.

Advanced Modeling Techniques: Explore advanced ML algorithms and methodologies, such as deep learning architectures or ensemble methods, to capture complex patterns and relationships within the data more effectively.

Data Augmentation: Augment the existing dataset with synthetic data generated through techniques like data synthesis or data simulation. This can help to address imbalances in the dataset and improve model generalization.

Improved Data Quality: Implement robust data preprocessing techniques to handle outliers, missing values, and inconsistencies in the data. Utilize quality assurance measures to ensure the accuracy and reliability of IoT sensor data.

Model Optimization and Hyperparameter Tuning: Fine-tune model parameters and hyperparameters using techniques like grid search or Bayesian optimization to optimize model performance and enhance predictive accuracy.

Cross-Validation and Model Evaluation: Employ rigorous cross-validation strategies to assess model performance and generalization ability. Validate models on diverse datasets and evaluate them using appropriate performance metrics to ensure reliable predictions.

Integration of Domain Expertise: Collaborate closely with domain experts, including dairy scientists and veterinarians, to incorporate domain knowledge into the modeling process. This can help to refine feature selection, interpret model results, and improve the overall accuracy of milk quality prediction.

Continuous Monitoring and Feedback: Establish mechanisms for continuous monitoring of model performance in real-world settings. Incorporate feedback loops to iteratively improve model accuracy based on new data and evolving requirements.

7. RESULTS & DISCUSSION

In this study, we utilized machine learning techniques to classify milk samples and detect the presence of adulterants. Various classification methods commonly employed in the food industry, including Random Forest, Neural Networks, and k-nearest Neighbor, were applied. The technologies and statistical methodologies employed to assess the quality parameters in milk samples are outlined below. The dataset was trained using various machine-learning models via sci-kit-learn, a comprehensive Python library for machine learning. We employed three models: Logistic Regression, Random Forest Classifier, and additional methods. The dataset underwent partitioning into training and testing subsets utilizing the train-test-split function. Following this, feature standardization was executed using Standard Scaler to guarantee consistency in scale across variables. Logistic Regression and Random Forest models were then instantiated, trained on the training data, and utilized for predicting the target variable within the testing set. Subsequently, model performance was assessed using the accuracy-score function to evaluate the effectiveness of each model on unseen data. This coding framework demonstrates a standard workflow for training, evaluating, and comparing different machine learning models, highlighting the versatility and user-friendly nature of scikit-learn for such tasks.

7.1. Logistic Regression:

Code

This is for the test dataset

model = LogisticRegression()

model.fit(X train scaled,y train)

X test predicts lr =model.predict(X test)

This is for the train dataset

model = LogisticRegression()

model.fit(X test scaled,y test)

X train predicts lr =model.predict(X train)

print('The accuracy for Logistic Regression model is (Train Dataset) ', metrics.accuracy score(X train predict lr,y train))

The confusion matrix for logistic regression shows the following values:

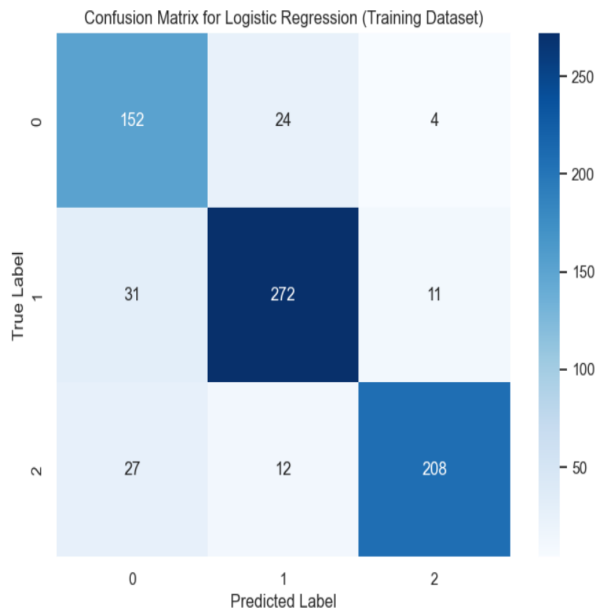

Figure 7.1.1: Confusion matrix for Logistic Regression(Train)

True Labels (0, 1, 2): These are the actual classes of the data points.

0: Represents the first class (low).

1: Represents the second class (medium).

2: Represents the third class (high).

Predicted Labels (0, 1, 2): These are the classes predicted by the model.

0: The model predicts the data point belongs to Class'low'.

1: The model predicts the data point belongs to class 'medium'.

2: The model predicts the data point belongs to class 'high'.

For Class 'low':

True Positives (TP): 152

False Positives (FP): 31 + 27 = 58

False Negatives (FN): 24 + 4 = 28

Misclassified = False Positives (FP) + False Negatives (FN) = 58 + 28 = 86

Correctly Classified = True Positives (TP) = 152

Quality Ratio = TP / (TP + FP + FN) = 152 / (152 + 58 + 28) = 0.637

For Class 'medium':

True Positives (TP): 272

False Positives (FP): 24 + 12 = 36

False Negatives (FN): 31 + 11 = 42

Misclassified = False Positives (FP) + False Negatives (FN) = 36 + 42 = 78

Correctly Classified = True Positives (TP) = 272

Quality Ratio = TP / (TP + FP + FN) = 272 / (272 + 36 + 42) = 0.769

For Class 'high':

True Positives (TP): 208

False Positives (FP): 4 + 11 = 15

False Negatives (FN): 27 + 12 = 39

Misclassified = False Positives (FP) + False Negatives (FN) = 15 + 39 = 54

Correctly Classified = True Positives (TP) = 208

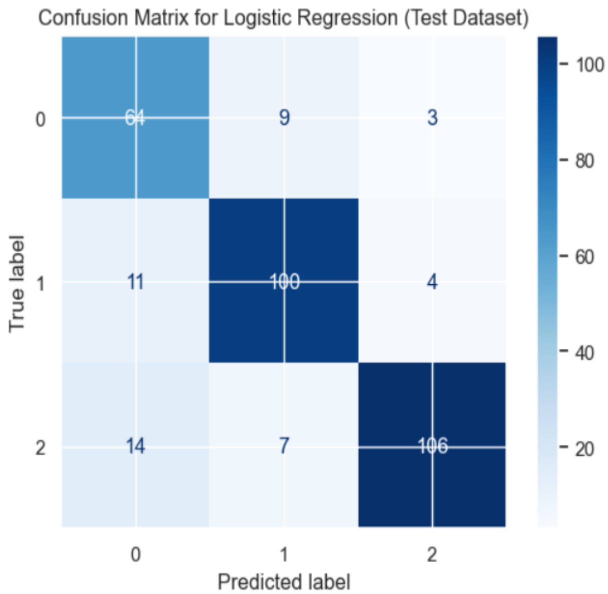

Figure 7.1.2: Confusion matrix for Logistic Regression(Test)

The confusion matrix for the logistic regression test data reveals the following values:

For Class 'low':

True Positives (TP): 64

False Positives (FP): 11 + 14 = 25

False Negatives (FN): 9 + 3 = 12

Misclassified = False Positives (FP) + False Negatives (FN) = 25 + 12 = 37

Correctly Classified = True Positives (TP) = 64

For Class 'medium':

True Positives (TP): 100

False Positives (FP): 9 + 7 = 16

False Negatives (FN): 11 + 4 = 15

Misclassified = False Positives (FP) + False Negatives (FN) = 16 + 15 = 31

Correctly Classified = True Positives (TP) = 100

For Class 'high': True Positives (TP): 106

False Positives (FP): 3 + 4 = 7

False Negatives (FN): 14 + 7 = 21

Misclassified = False Positives (FP) + False Negatives (FN) = 7 + 21 = 28

Correctly Classified = True Positives (TP) = 106

7.2. Random Forest:

Code:

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy score
model = RandomForestClassifier(n estimators=100)
model.fit(X train, y train)
X test predict rfc = model.predict(X test)
print('The accuracy of the Random Forests model is (Test Dataset)',
accuracy score(X test predict rfc, y test))
model train = RandomForestClassifier(n estimators=100)
model train.fit(X test, y test)
X train predict rfc = model train.predict(X train)
print('The accuracy of the Random Forests model is (Train Dataset)',
accuracy score(X train predict rfc, y train))
```

The confusion matrix for random forest shows the following values:

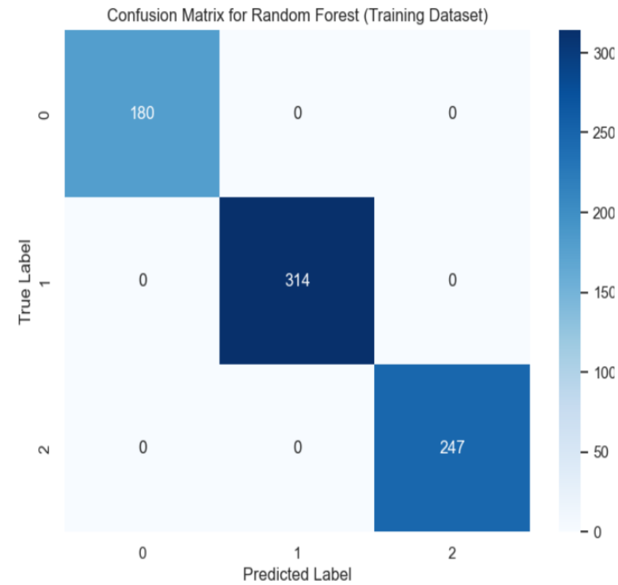

Figure 7.2.1: Confusion matrix for Random Forest(Train)

For Class ’low’:

True Positives (TP): 180

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 180

For Class ’medium’:

True Positives (TP): 314

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 314

For Class 'high':

True Positives (TP): 247

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 247

The confusion matrix for the Random Forest test data presents the following values:

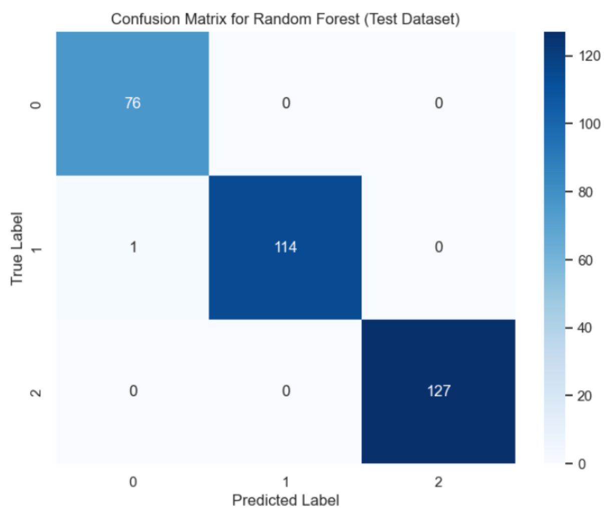

Figure 7.2.2: Confusion matrix for Random Forest(Test)

For Class ’low’:

True Positives (TP): 76

False Positives (FP): 1

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 1 + 0 = 1

Correctly Classified = True Positives (TP) = 76

For Class ’medium’:

True Positives (TP): 114

False Positives (FP): 0

False Negatives (FN): 1

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 1 = 1

Correctly Classified = True Positives (TP) = 114

For Class ’high’:

True Positives (TP): 127

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 127

7.3. Decision Tree Classifier:

Code:

```
model = DecisionTreeClassifier()
model.fit(X train,y train)
X test predict dtc = model.predict(X test)
print('The accuracy for Decision Tree Classifier model is (Test Dataset) ',
metrics.accuracy score(X test predict dtc,y test))
model = RandomForestClassifier(n estimators=100)
model.fit(X test,y test)
X train predict =model.predict(X train)
print('The accuracy for Decision Tree Classifier model is (Train Dataset)
',metrics.accuracy score(X train predict,y train))
```

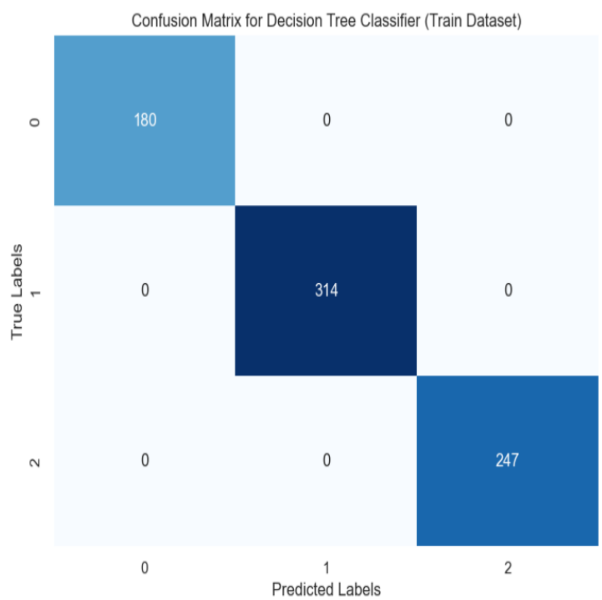

Figure 7.3.1: Confusion matrix for the Decision Tree Classifier(Train)

The confusion matrix for the Decision Tree Classifier (Train Dataset):

For Class 'low':

True Positives (TP): 314

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 314

For Class 'medium':

True Positives (TP): 247

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 247

For Class 'high':

True Positives (TP): 180

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 180

The confusion matrix generated by the Decision Tree Classifier (Test) exhibits the following values:

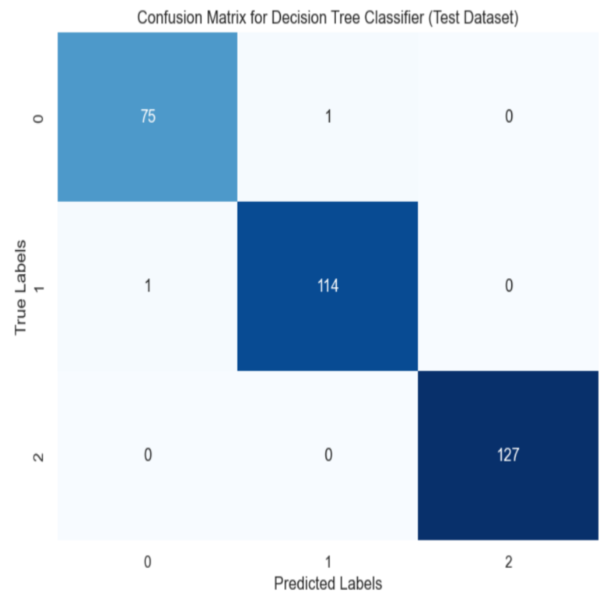

Figure 7.3.2: Confusion matrix for the Decision Tree Classifier(Test)

For Class 'low':

True Positives (TP): 75

False Positives (FP): 1

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 1 + 0 = 1

Correctly Classified = True Positives (TP) = 75

For Class 'medium':

True Positives (TP): 114

False Positives (FP): 1

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 1 + 0 = 1

Correctly Classified = True Positives (TP) = 114

For Class ’high’:

True Positives (TP): 127

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 127

7.4. K-Nearest Neighbour(KNN):

Code:

model = KNeighborsClassifier()

model.fit(X train,y train)

X test predict knn = model.predict(X test)

print(’The accuracy for KNN classifier model is (Test Dataset)’,

metrics.accuracy score(X test predict knn,y test)) model = KNeighborsClassifier()

model.fit(X test,y test)

X train predict knn =model.predict(X train)

print(’The accuracy for KNN classifier model is (Train Dataset)

’,metrics.accuracy score(X train predict knn,y train))

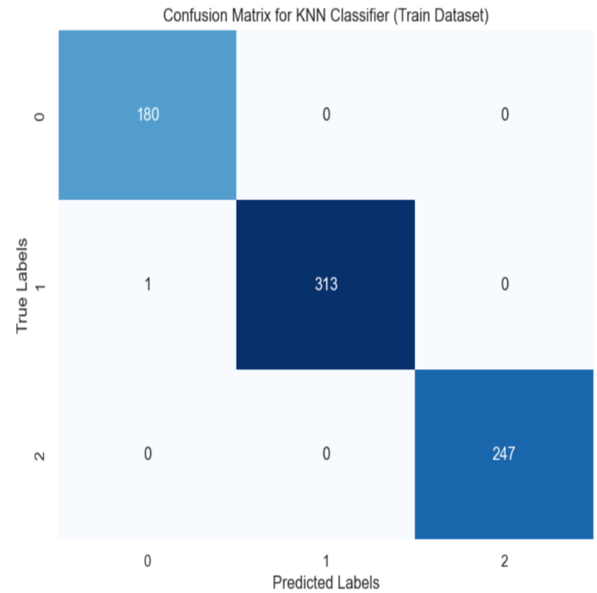

Figure 7.4.1: Confusion matrix for KNN classifier(Train)

The confusion matrix for KNN Classifier (Train Dataset):

For Class 'low':

True Positives (TP): 180

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 180

For Class 'medium':

True Positives (TP): 313

False Positives (FP): 1

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 1 + 0 = 1

Correctly Classified = True Positives (TP) = 313

For Class 'high':

True Positives (TP): 247

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 247

The confusion matrix generated by the KNN Classifier using the Test Dataset is as follows:

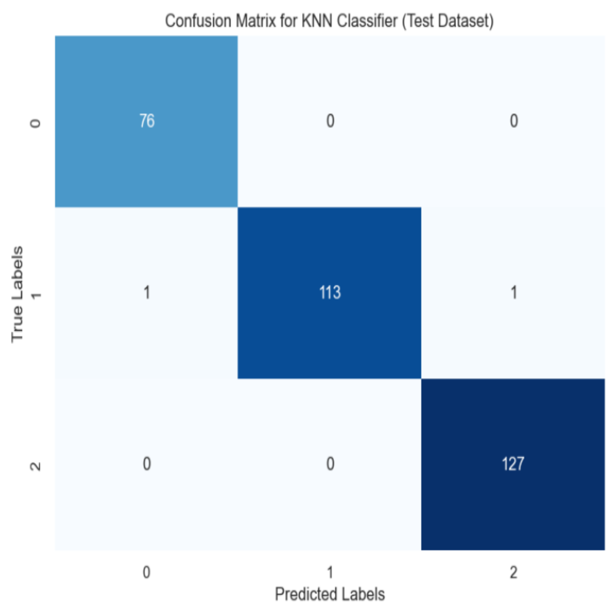

Figure 7.4.2: Confusion matrix for KNN classifier(Test)

For Class 'low':

True Positives (TP): 76

False Positives (FP): 1

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 1 + 0 = 1

Correctly Classified = True Positives (TP) = 76

For Class 'medium':

True Positives (TP): 113

False Positives (FP): 0

False Negatives (FN): 1

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 1 = 1

Correctly Classified = True Positives (TP) = 113

For Class 'high':

True Positives (TP): 127

False Positives (FP): 0

False Negatives (FN): 0

Misclassified = False Positives (FP) + False Negatives (FN) = 0 + 0 = 0

Correctly Classified = True Positives (TP) = 127

The data is trained with different models as shown below and obtained the following accuracies.

Analysis of used methods:

Method	Accuracy
Logistic Regression	85.2%
Random Forest	99.8%
Decision Tree Classifier	99.3%
K-nearest neighbor (KNN)	99.3%

Figure 7.4.3: Performance Analysis of used methods

7.5. Webapge Results

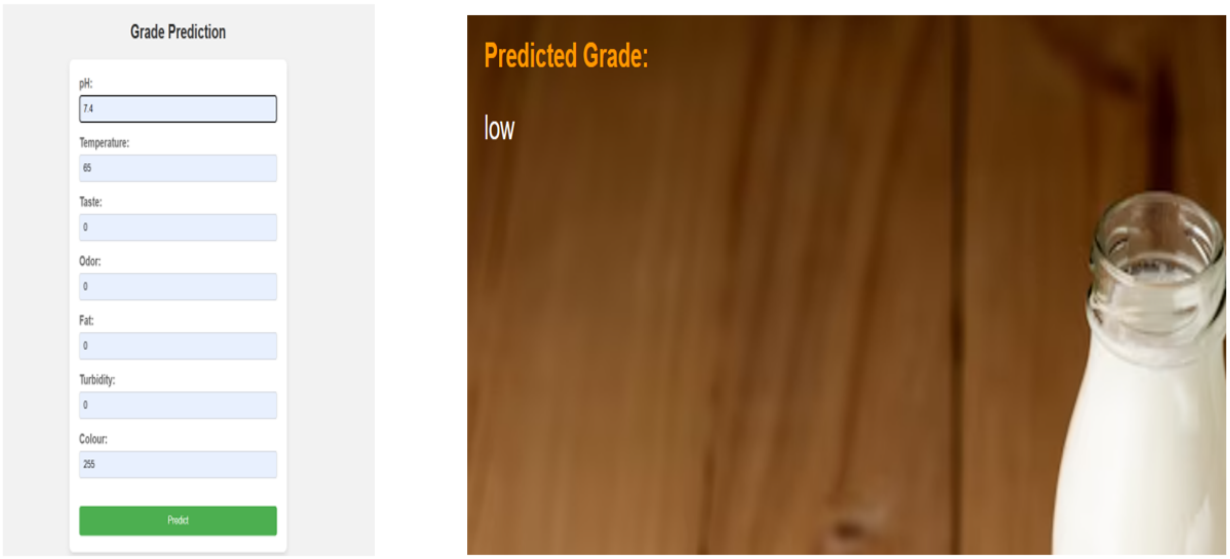

Figure 7.5.1: Low Quality

The image includes categories such as “Temperature,” “Odor,” “Fat,” “Turbidity,” and “Colour”. Each category has corresponding scores. The overall quality of the milk is low.

This assessment helps ensure that the milk meets established standards for safety, consistency, and consumer satisfaction before it is processed or distributed.

The categories listed on the form are crucial for determining the overall quality of milk.

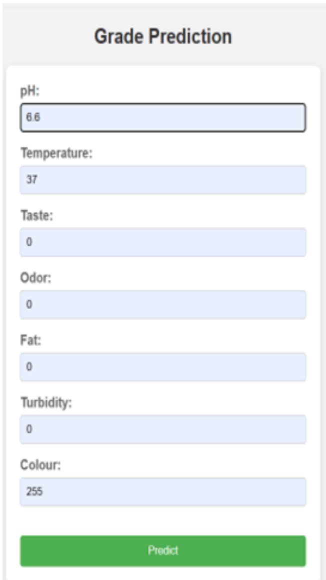

Figure 7.5.2: Medium Quality

The image includes categories such as "Temperature," "Odor," "Fat," "Turbidity," and "Colour". Each category has corresponding scores. The overall quality of the milk is medium.

This assessment helps ensure that the milk meets established standards for safety, consistency, and consumer satisfaction before it is processed or distributed.

The categories listed on the form are crucial for determining the overall quality of milk.

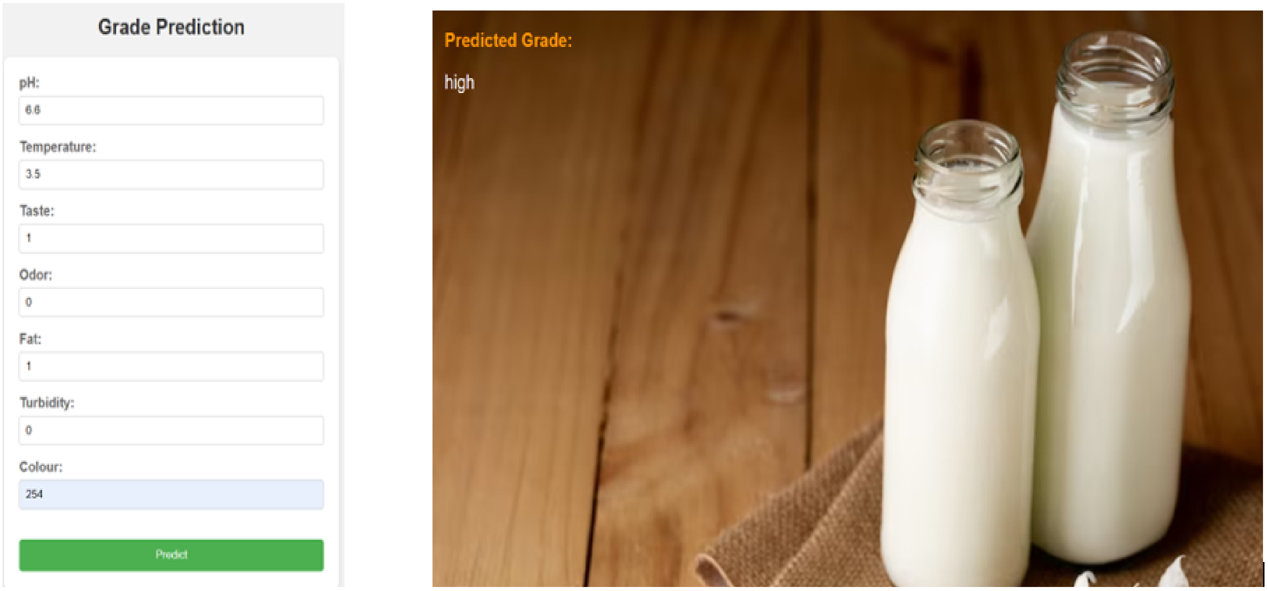

Figure 7.5.3: High Quality

The image includes categories such as “Temperature,” “Odor,” “Fat,” “Turbidity,” and “Colour”. Each category has corresponding scores. The overall quality of the milk is high.

This assessment helps ensure that the milk meets established standards for safety, consistency, and consumer satisfaction before it is processed or distributed.

The categories listed on the form are crucial for determining the overall quality of milk.

8. CONCLUSIONS & FUTURE SCOPE

The milk quality monitoring system, integrating IoT and ML technologies, serves as a pivotal tool in safeguarding the quality and safety of dairy products. By harnessing the power of Internet of Things (IoT) devices such as sensors and actuators, coupled with machine learning (ML) algorithms, this system enables real-time monitoring and analysis of key parameters in milk production and processing. It facilitates the continuous evaluation of factors like temperature, acidity, presence of contaminants, and overall composition, ensuring adherence to stringent quality standards throughout the production chain.

At the heart of this system lies a sophisticated machine learning model trained to interpret data collected from IoT sensors and make informed decisions regarding milk quality. Through the utilization of ML techniques such as classification, regression, and anomaly detection, the model can accurately identify deviations from optimal quality parameters and promptly alert stakeholders to potential issues. Moreover, the system can adapt and improve over time through continuous learning from new data, enhancing its ability to predict and prevent quality-related incidents.

The implementation of a milk quality monitoring system empowered by IoT and ML harbors significant implications for various stakeholders across the dairy industry. From farmers striving to optimize herd health and milk production to processors committed to delivering safe and nutritious dairy products, this technology offers unprecedented insights and control. Furthermore, consumers stand to benefit from increased transparency and trust in the dairy supply chain, as they gain access to milk products that meet rigorous quality standards. By harnessing the synergies between IoT and ML, this system represents a transformative step towards ensuring the integrity and sustainability of milk production while safeguarding public health.

Furthermore, the integration of IoT and ML technologies in milk quality monitoring not only enhances the efficiency and reliability of the production process but also fosters innovation in the dairy industry. By harnessing real-time data insights, stakeholders can implement preventive measures and optimize resource utilization, leading to improved productivity and sustainability. Additionally, the continuous monitoring and analysis provided by the system offer opportunities for research and development, driving advancements in dairy farming practices and product quality.

Moreover, as the system evolves, it has the potential to integrate with other emerging technologies such as blockchain, enabling end-to-end traceability and further enhancing transparency in the dairy supply chain. This integration can facilitate seamless tracking of milk from farm to table, providing consumers with detailed information about the origin, handling, and quality of the products they purchase.

REFERENCES

[1] H. Abdu and M. H. M. Noor, “A survey on waste detection and classification using deep learning,” IEEE Access, vol. 10, pp. 128151–128165, 2022.

[2] S. Shahab, M. Anjum, and M. S. Umar, “Deep learning applications in solid waste management: A deep literature review,” Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 3, pp. 381–395, 2022.

[3] Yadav, Sachin, et al. ”Waste classification and segregation: Machine learning and iot approach.”2021 2nd international conference on intelligent engineering and management (ICIEM). IEEE, 2021.

[4] N. Sarker, S. Chaki, A. Das, and M. S. A. Forhad, “Illegal trash thrower detection based on hogsvm for a real-time monitoring system,” pp. 483–487, 2021.

[5] X. Xu, X. Qi, and X. Diao, “Reach on waste classification and identification by transfer learning and lightweight neural network,” 2020.

[6] B. D. Carolis, F. Ladogana, and N. Macchiarulo, “YOLO TrashNet: Garbage detection in video streams,” in Proc. IEEE Conf. Evolving Adapt. Intell. Syst. (EAIS), May 2020.

[7] M. Fulton, J. Hong, M. J. Islam, and J. Sattar, “Detection of marine litter using deep visual detection models,” in Proc. Int. Conf. Robot. Autom. (ICRA), May 2019

[8] M. Pamintuan, S. M. Mantiquilla, H. Reyes, and M. J. Samonte, “i-bin: An intelligent trash bin for automatic waste segregation and monitoring system,” pp. 1–5, 2019.

[9] C. Bircano glu, M. Atay, F. Bes er, O. Genc , and M. A. Kızrak, “Recyclenet: Intelligent waste sorting using deep neural networks,” pp. 1–7, 2018.

[10] S. Sudha, M. Vidhyalakshmi, K. Pavithra, K. Sangeetha, and V. Swaathi, “An automatic classification method for environment: Friendly waste segregation using deep learning,” pp. 65–70, 2016.