iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VII **Month of publication:** July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84486

Smart Milk Adulteration & Dairy Fraud Detector: An IoT and Machine Learning Based Approach

S Sri Lakshmi[1], CH Charan[2], G Naga Bharath Raj[3], K Bharath Chary[4], N Rohin Sai[5]
Department of Computer Science and Engineering, Vidya Jyothi Institute of Technology, Hyderabad, India

Abstract: Milk adulteration is a critical public health concern, particularly in developing countries, where harmful substances such as water, urea, and detergent are commonly added to increase quantity and profit. Conventional detection methods rely on laboratory-based chemical analysis, which is time-consuming, expensive, and requires skilled personnel, creating a strong need for an automated, cost-effective, and portable solution for instant milk quality assessment. This paper presents a Smart Milk Adulteration & Dairy Fraud Detection System that integrates Internet of Things (IoT) and Machine Learning (ML) technologies to enable real-time detection. The system uses an Arduino UNO combined with pH, temperature (DHT11), and conductivity sensors to collect milk sample data, which is transmitted through an ESP8266 WiFi module to a Node.js backend and further processed by a Python Flask-based ML service. A Random Forest algorithm classifies milk into four categories -- Pure, Watered, Detergent, and Urea adulterated -- with results displayed instantly on an LCD screen and a React.js web dashboard, while all readings are stored in a MongoDB database. The proposed system achieves a cross-validation F1-macro score of 99.25%, with a total hardware cost under ₹4,000, offering a reliable, low-cost, and scalable solution for real-time milk adulteration detection suitable for dairy farms, milk collection centers, and food safety monitoring.
Keywords: Milk Adulteration, Internet of Things, Machine Learning, Random Forest, Arduino UNO, Real-Time Monitoring, Food Safety

I. INTRODUCTION

Milk is one of the most essential and widely consumed food products across the world due to its rich nutritional value, including proteins, vitamins, and minerals. In countries like India, the dairy industry forms a significant part of the economy. However, increasing demand has led to widespread adulteration practices, where harmful substances such as water, urea, and detergent are mixed into milk to increase volume and profit, posing serious health risks to consumers.

Traditional detection methods rely on laboratory-based chemical testing, which requires specialized equipment, trained personnel, and considerable time, making them unsuitable for on-site or real-time detection, especially in rural areas and small-scale dairy collection centers. This gap highlights the need for a modern, automated system that can provide quick and reliable results without depending on laboratory infrastructure.

Advancements in IoT and Machine Learning have opened new possibilities for real-time monitoring and intelligent decision-making. IoT enables the collection of real-time sensor data, while ML algorithms analyze this data to classify outcomes accurately. In this work, a Smart Milk Adulteration & Dairy Fraud Detection System is developed using an Arduino UNO for sensor data collection, an ESP8266 WiFi module for transmission, and a Random Forest ML model for classification, with results displayed instantly on an LCD screen and a web dashboard.

II. LITERATURE SURVEY

Milk adulteration detection has been a critical research area due to its direct impact on public health and food safety. Traditional laboratory-based chemical analysis techniques, although accurate, require sophisticated equipment, skilled personnel, and significant processing time, making them unsuitable for real-time or on-site applications, particularly in rural and large-scale dairy supply chains.

Recent advancements in IoT have enabled smart systems for food quality monitoring using pH, conductivity, and temperature sensors, which provide quicker results than conventional methods but often lack integration with intelligent data analysis techniques. In parallel, Machine Learning algorithms such as Decision Trees, Support Vector Machines, and Random Forest have been applied to classify milk samples based on sensor data, with Random Forest reported to achieve higher accuracy owing to its ability to handle multiple input parameters and reduce overfitting; however, most existing models are validated only in controlled environments.

Portable testing devices integrated with microcontrollers such as Arduino UNO have also been explored to provide low-cost, accessible solutions for farmers and dairy industries, though they often lack real-time visualization and centralized storage. Modern systems increasingly incorporate web technologies and cloud databases such as React.js and MongoDB to enable real-time monitoring, yet a gap remains in integrating IoT sensing, ML classification, and user-friendly dashboards into a single unified system -- a gap this work directly addresses.

III. PROPOSED SYSTEM

A. Existing System and Limitations

Existing milk adulteration testing relies on laboratory methods such as the Gerber method, the formalin spot test, the starch-iodine test, and the lactometer test. These require trained technicians and specialized equipment, take hours to days to yield results, cost ₹500--₹2,000 per test, are not portable, and provide no historical logging or trend analysis.

B. Proposed System Overview

The proposed system addresses these limitations by delivering results in under 10 seconds using IoT sensors and ML classification, at a total hardware cost under ₹4,000. It requires no trained technician, offers dual result display on an LCD (standalone) and a React.js dashboard (networked), automatically logs all readings to MongoDB with timestamps, and achieves 99.25% F1-macro accuracy across four adulteration classes. A "Test Now" button on the dashboard additionally enables on-demand testing.

IV. SYSTEM DESIGN

A. System Architecture

The system architecture consists of four interconnected layers: the hardware sensor layer (Arduino UNO + ESP8266), the backend server layer (Node.js + Express.js + MongoDB), the ML service layer (Python Flask + Random Forest), and the frontend dashboard layer (React.js), as shown in Fig. 1.

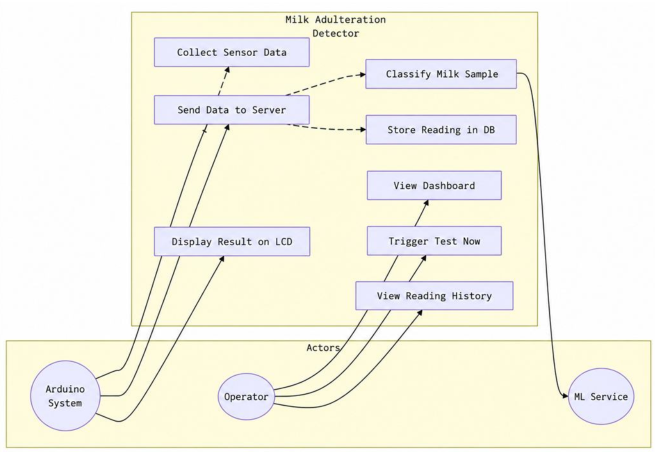


Fig. 1 System architecture

B. Hardware Requirements

- Microcontroller: Arduino UNO (ATmega328P)
- WiFi Module: ESP8266 (ESP-01) for wireless communication
- pH Sensor: Analog pH probe module
- Conductivity Sensor: for detecting urea and detergent adulterants
- Temperature Sensor: DHT11
- Display Unit: 16x2 LCD (JHD 162A)

C. Software Requirements

The backend uses Node.js v18+ with Express.js, Mongoose, and Axios; the ML service uses Python Flask with scikit-learn and NumPy; the frontend dashboard is built with React.js and Vite; and MongoDB Community Edition serves as the database. All components are free and open-source.

V. METHODOLOGY AND IMPLEMENTATION

A. Firmware (Arduino)

The Arduino sketch initializes the LCD, DHT11 sensor, and a SoftwareSerial connection to the ESP8266. In each cycle it reads the pH and conductivity sensors and the DHT11 temperature, constructs a JSON payload, and transmits it via HTTP POST to the Node.js backend using AT commands. It then performs an HTTP GET to fetch the classification result and displays it on the LCD. The pH value is computed using $ph = 7.0 + (676 - rawAnalogValue) \times 0.0178$, calibrated against a distilled-water reference.

B. Backend Server

The backend, built on Node.js and Express.js, exposes REST endpoints for receiving sensor data, forwarding it to the ML service via Axios, storing results in MongoDB via Mongoose, and serving the latest result and trigger flag to the Arduino and dashboard. CORS middleware allows cross-origin access from the React dashboard.

C. Machine Learning Service

A Flask application serves a Random Forest classifier (300 estimators, max depth 12, balanced class weights) trained on a synthetic dataset of 400 samples spanning Pure, Watered, Detergent, and Urea classes, generated using scientifically validated parameter ranges. The /predict endpoint returns the predicted class and confidence percentage, achieving a cross-validation F1-macro score of 99.25% (±1.51%).

D. Frontend Dashboard

The React.js dashboard, built with Vite, polls the backend every 10 seconds to display the latest reading -- pH, temperature, conductivity, result, and confidence -- in a card layout, alongside a history table of past readings. A "Test Now" button triggers an on-demand reading cycle on the Arduino.

VI. RESULTS AND DISCUSSION

The system was evaluated using functional, integration, and unit testing. Functional testing confirmed correct operation of the "Test Now" trigger, automatic dashboard refresh, and the reading-history table. Integration testing verified end-to-end communication between the Arduino, backend, Flask ML service, and MongoDB. Unit testing confirmed each sensor operated within its expected range -- the pH sensor read 6.8--7.2 in distilled water and the conductivity sensor read 400--600 µS/cm in tap water.

The Random Forest classifier achieved a cross-validation F1-macro score of 99.25% (±1.51%), exceeding the 90% target accuracy. Results are returned within seconds of sensor acquisition and displayed on both the LCD and web dashboard, with all readings logged for historical trend analysis. The total hardware cost of under ₹4,000 makes the system economically viable for deployment at dairy collection centers and by local vendors.

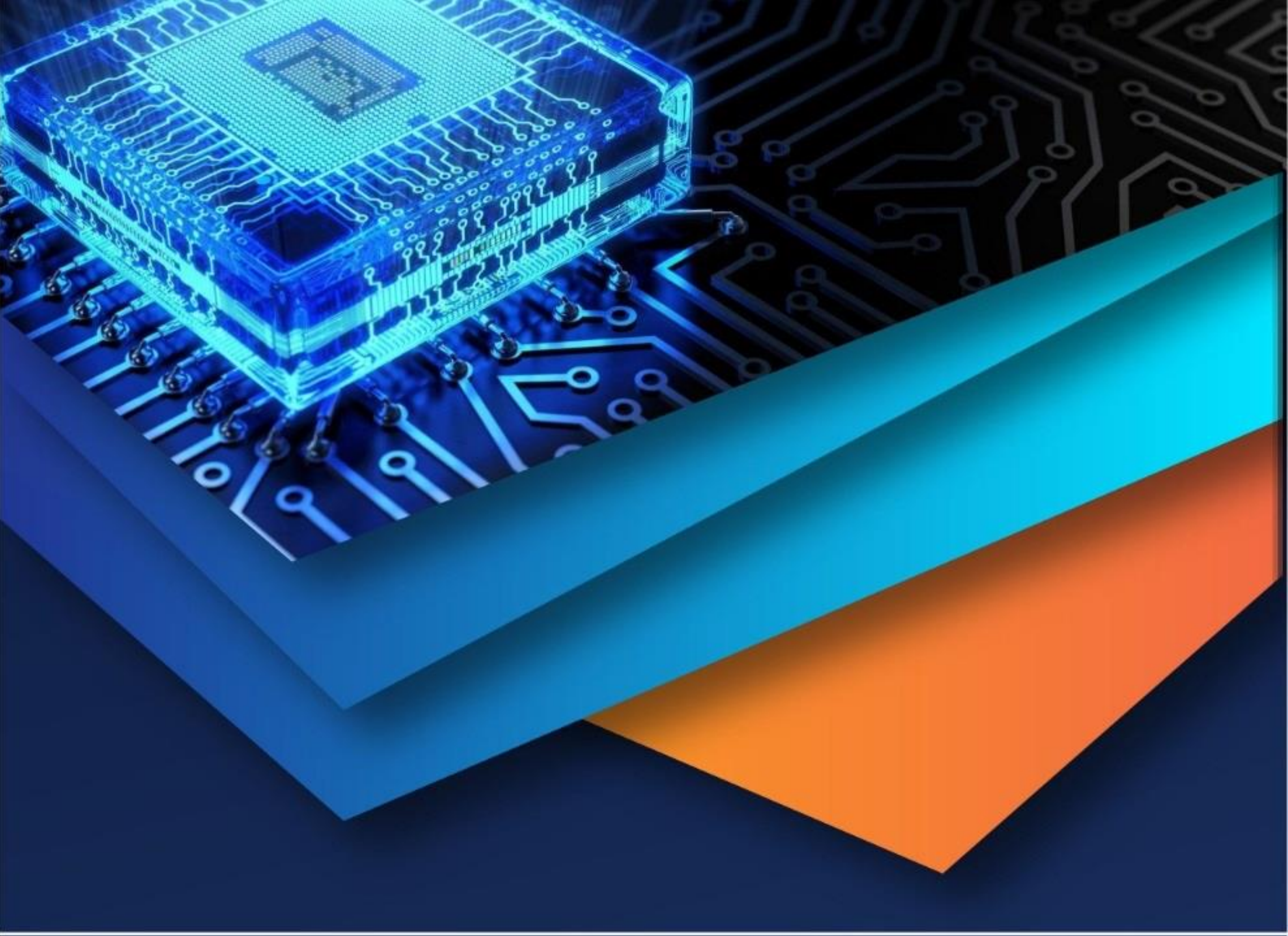
VII. CONCLUSION AND FUTURE SCOPE

This paper presented a Smart Milk Adulteration & Dairy Fraud Detector that combines low-cost IoT hardware with machine learning for real-time dairy fraud detection. The system integrates an Arduino UNO with pH, conductivity, and temperature sensors, a Node.js/Express.js backend, a Flask-based Random Forest ML service, a MongoDB database, and a React.js dashboard, achieving 99.25% F1-macro accuracy at a hardware cost under ₹4,000.

Future work includes retraining the model on real laboratory-collected samples across diverse seasons and regions, adding spoiled/curdled milk and additional adulterant detection (starch, salt, formalin), integrating WhatsApp/Telegram alerts, adding a standalone physical trigger button, and deploying the system to the cloud (Render and Vercel) with a companion mobile application for field inspectors.

REFERENCES

[1] Sharma, R., & Rajput, Y. S., "Detection of adulteration in milk -- A review," Indian Journal of Dairy Science, vol. 67, no. 3, pp. 1--10, 2014.
[2] Patel, S., & Sinha, A., "IoT based milk quality monitoring system using sensors," Int. J. Eng. Research & Technology (IJERT), vol. 8, no. 5, pp. 112--116, 2019.
[3] Naila, A., Flint, S., Fletcher, G., Bremer, P., & Meerdink, G., "Chemical indicators of hydrolytic rancidity in milk," Journal of Food Science, vol. 75, no. 6, pp. 105--111, 2010.
[4] Breiman, L., "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5--32, 2001.
[5] Pedregosa, F., et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825--2830, 2011.
[6] Arduino, "Arduino UNO Reference," 2024. [Online]. Available: https://docs.arduino.cc
[7] Espressif Systems, "ESP8266 AT Instruction Set," 2024. [Online]. Available: https://www.espressif.com
[8] MongoDB, "MongoDB Manual," 2024. [Online]. Available: https://www.mongodb.com/docs/manual
[9] React, "React Documentation," 2024. [Online]. Available: https://react.dev
[10] FSSAI, "Milk and Milk Products Standards," Food Safety and Standards Authority of India, 2023. [Online]. Available: https://www.fssai.gov.in



doi crossref

10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY