Milk quality prediction and yogurt fermentation analysis using Machine Learning

Prof. S. B. Ghawate[1], Shrihari Shinde[2], Jayesh Banda[3], Nisarg Wani[4], Aditya Nimse[5]
[1,2,3,4,5]*Computer Engineering, SAE, Kondhwa*

Abstract: **During the development of innovative products, consumer preferences are the essential factors for yogurt producers to improve their market share. A high-performance prediction method will be beneficial to understanding the intrinsic relevance between preferences and sensory attributes. In this study, a novel deep learning method is proposed that uses an autoencoder to extract product features from the sensory attributes scored by experts, and the sensory features bought are regressed on consumer preferences with support vector machine analysis. Ensuring Milk quality is crucial for producing high-quality dairy products, such as yogurt. This study aims to develop a comprehensive approach for predicting Milk quality and analyzing yogurt fermentation. The proposed method combines machine learning techniques for Milk quality prediction and advanced bioprocess analysis for yogurt fermentation.**

Keywords: yogurt; Model-train; consumer preference; autoencoder; support vector Machine.

I. INTRODUCTION

Machine learning (ML) techniques have become a pivotal force in the food industry, particularly in the quest to uphold quality control and food safety standards. One of the significant applications in this domain is the correct differentiation between Milk and yogurt, both essential products in the dairy industry. Achieving this distinction is not just a matter of convenience; it is crucial for safeguarding product quality, preventing contamination, and ensuring authenticity in the dairy sector. To address this challenge and harness the potential of ML, a multi-faceted approach is adopted.[3]

The first step in creating an ML model for distinguishing between Milk and yogurt is the compilation of a comprehensive dataset. This dataset typically includes images, spectroscopic data, or chemical properties of both Milk and yogurt. Gathering a diverse range of data is essential to ensure the model's accuracy in various real-world scenarios. Once collected, this data undergoes rigorous pre-processing. This step involves noise removal, standardization of data formats, and the establishment of data consistency across the entire dataset.

With a clean and well-organized dataset at hand, various ML algorithms come into play. These include support vector machines (SVM), random forests, and deep learning models like Convolutional Neural Networks (CNNs).[2] These algorithms are then trained on the pre-processed dataset to develop a robust model capable of distinguishing between Milk and yogurt with high accuracy.

The motivation behind applying machine learning to predict Milk quality and perfect yogurt fermentation processes is far-reaching. It transcends merely separating the two dairy products. The core aims extend to enhancing the dairy industry's efficiency, ensuring product quality, and driving innovation in an ever-evolving landscape.[7] By using data-driven insights, machine learning opens the doors to more efficient, sustainable, and competitive practices in dairy production.

To achieve these aims successfully, a combination of factors is needed. It's not merely about implementing machine learning algorithms but also about tapping into domain expertise and meticulous data engineering. These elements collectively play a crucial role in achieving the goal of accurately distinguishing between Milk and yogurt.

The practical application of such a system is significant.[5] It can run in real time, whether on production lines, at dairy farms, or within retail environments. This ensures prompt intervention in cases of mislabelling or contamination, thereby safeguarding consumer safety and maintaining product authenticity. In essence, the dairy industry stands to benefit immensely from the integration of machine learning techniques, with the potential to elevate product quality, streamline processes, and foster innovation on the horizon.

II. KEY ASPECTS OF MILK & YOGURT FERMENTATION DETECTION PROJECT

Key aspects of a Milk and yogurt Fermentation Detection Project, as outlined in the provided information, include:

ASPECT 1: Significance in the Food Industry Milk and yogurt detection using machine learning is essential in the food industry, particularly for quality control and food safety.

ASPECT 2: Differentiation Challenge The primary challenge is accurately distinguishing between milk and yogurt, which is crucial for quality control, contamination prevention, and ensuring product authenticity.

ASPECT 3: Data Collection A diverse and comprehensive dataset of images, spectroscopic data, or chemical properties of both milk and yogurt is collected.

ASPECT 4: Data Preprocessing The collected data is pre-processed to remove noise, standardize data format, and ensure consistency across the dataset.

ASPECT 5: Machine Learning Algorithms Various machine learning algorithms are employed, including support vector machines (SVM), random forests, and deep learning models like Convolutional Neural Networks (CNNs).

ASPECT 6: Motivation The motivation behind this project is to improve product quality, perfect processes, ensure consumer safety, and drive innovation in the dairy industry through data-driven insights.

ASPECT 7: Efficiency and Sustainability The project aims to lead to more efficient, sustainable, and competitive dairy production practices.

ASPECT 8: Quality Prediction One aim is to develop a machine learning model to predict the quality of milk based on various parameters.

ASPECT 9: Fermentation Analysis The other aim is to employ machine learning techniques to analyze and perfect the yogurt fermentation process.

ASPECT 10: Multidisciplinary Approach Success in these aims requires a combination of domain ability, data engineering, and advanced machine learning techniques.

ASPECT 11: Real-time Detection Implementing a system capable of detecting the presence of milk and yogurt in real-time, whether in production lines, dairy farms, or retail environments, is a key aspect of the project.

Overall, this project combines domain knowledge, data processing, and advanced machine learning to address the challenge of accurately distinguishing between milk quality and yogurt quality by considering various parameters, to enhance efficiency and product quality in the dairy industry.

III. LITERATURE REVIEW

Currently, the prediction of Milk yield can help pasture managers coordinate production and transportation planning for a farm on time. The algorithm introduces the genetic algorithm (GA) into the long short-term memory (LSTM) algorithm structure, which considers the time sequence and correlations between the above input variables. The experiment proves that the GA-LSTM algorithm is more correct and stable than the traditional LSTM algorithm in predicting daily Milk yields.[1]

This paper presents the advanced control theory's original utilization to realize a system that controls the fermentation process in batch bioreactors. Proper fermentation control is essential for quality fermentation products and the economical operation of bioreactors. Batch bioreactors are extremely popular due to their simple construction. However, this simplicity presents limitations in implementing control systems that would ensure a controlled fermentation process. Batch bioreactors do not allow the inflow/outflow of substances during operation. Therefore, we have developed a control system based on a stirrer drive instead of material flow.[2]

Milk is a highly perishable product, whose quality degrades while moving downstream in an imperfect cold dairy supply chain. Existing literature adopts a reactive approach for evaluating and preventing Milk with a high microbial index from moving further downstream in a dairy supply chain. In this paper, we argue that such an approach is not the best response if the intention is to maximize Milk life in terms of quality. We propose an initiative-taking approach that

checks the metrics of the temperature and the level that are the building blocks of microorganisms in Milk.[3]
In this study, a novel deep learning method is proposed that uses an autoencoder to extract product features from the sensory attributes scored by experts, and the sensory features bought are regressed on consumer preferences with support vector machine analysis. Model performance analysis, hedonic contour mapping, and feature clustering were implemented to confirm the overall learning process. The results showed that the deep learning model can vouch for an acceptable level of accuracy, and the hedonic mapping reflected could supply immense help for producers' product design or modification.[5]
Quantitative descriptive analysis was used to describe the key attributes of nine ultra-pasteurized (UP) Milk products of various fat levels, including two lactose-reduced products, from two dairy plants. Principal components analysis found four significant principal components that accounted for 87.6 of the variances in the sensory attribute data. Principal part scores showed that the location of each UP Milk along each of the four scales primarily corresponded to cooked, drying/lingering, sweet, and bitter attributes.[5]
Numerous statistical machine learning methods suitable for application to highly correlated features, such as those that exist for spectral data, could potentially improve prediction performance over the commonly used partial least squares approach. Milk samples from 622 individual cows with known detailed protein composition and technological trait data accompanied by mid-infrared spectra were available to assess the predictive ability of different regression and classification algorithms.[6]
In the dairy industry, machine learning techniques are being employed to predict and improve the quality of milk products. One case study focuses on using a random forest regression model to predict milk quality based on parameters like fat content, protein content, bacterial count, storage conditions, and processing methods. The model's success enables real-time quality assessment, allowing for prompt intervention and high-quality product production. Another case study concentrates on yogurt fermentation optimization, utilizing deep learning models and statistical techniques to analyze a dataset containing information about milk, bacterial strains, fermentation conditions, and sensory evaluation. The outcome is enhanced taste, texture, and consistency, leading to cost savings and quality improvement for dairy producers. Real-time adjustments during fermentation ensure product consistency and quality. These case studies highlight the practical application of machine learning in the dairy industry, enhancing efficiency, reducing waste, and ensuring product quality.[7]

IV. ALGORITHMS

1. Milk Quality Prediction*:*
a. *Support Vector Machines (SVM):* SVM is a popular algorithm for binary classification tasks like distinguishing between Milk and yogurt. It can be used to predict Milk quality based on various parameters.[1]
b. *Principal Component Analysis (PCA):* PCA can be applied for dimensionality reduction and visualization of Milk quality data.[3]
c. *K-Nearest Neighbours (K-NN):* K-NN is a simple and effective algorithm for classification tasks that can be used to predict Milk quality.[7]
2. Yogurt Fermentation Analysis:
a. *Neural Networks (CNN):* CNNs can capture temporal dependencies in yogurt fermentation data, making them suitable for analysing fermentation processes.[2]
b. *Long Short-Term Memory (LSTM):* LSTM, a type of RNN, can model long-term dependencies in time-series data, making it useful for tracking changes during yogurt fermentation.[5]
c. *Time Series Analysis:* Techniques like autoregressive integrated moving average (ARIMA) and seasonal decomposition of time series (STL) can be used for time-series data analysis in yogurt fermentation.[7]
d. Cluster Analysis: Clustering algorithms like K-Means and DBSCAN can help group similar yogurt fermentation patterns, aiding in optimization.[4]

3. Data Engineering and Preprocessing:
a. Data Cleaning: Techniques for noise removal, outlier detection, and managing missing data are crucial in preparing the dataset.
b. Feature Engineering: Create relevant features from the collected data, such as extracting spectral features or chemical properties.
c. Data Standardization: Normalize and standardize data to ensure consistency across the dataset.
4. Evaluation Metrics:

a. Use proper evaluation metrics such as accuracy, precision, recall, F1-score, and ROC AUC to assess the performance of your ML models.

V. SYSTEM ARCHIETECTURE

System Architecture for Machine Learning-Based Milk & Yogurt Fermentation detection

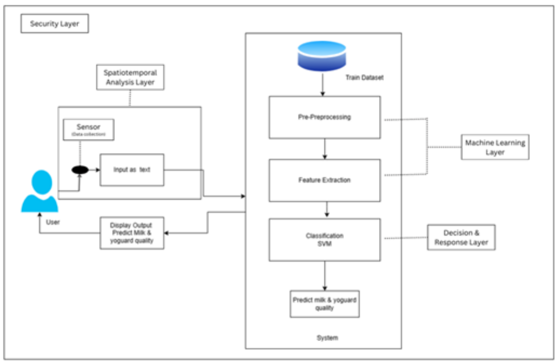

Figure 1: SYSTEM ARCHITECTURE DIAGRAM

1. Data Collection and Preprocessing:
- Gather diverse datasets for milk and yogurt.
- Clean, standardize, and format data, including noise removal and feature extraction.

2. Machine Learning Model Development:
- Predict Milk Quality and optimize Yogurt Fermentation

Using ML algorithms like SVM, Random Forests.

3. Real-Time Detection System:
- Implement real-time detection using cameras or sensors for classification.
- Deploy the system at various dairy industry points.

4. Integration and Deployment:
- Integrate ML models into production processes.
- Deploy real-time detection system on-premises cloud.

5. Feedback Loop:
- Continuously monitor and adapt the system to changing conditions.
- Incorporate insights from domain experts.

6. User Interface and Reporting:
- Provide a user-friendly interface for real-time interaction.
- Generate reports for decision-making and quality control.

7. Security and Authentication:
- Implement data security measures and user authentication.

Spatiotemporal Analysis Layer: A Spatiotemporal Analysis Layer is employed for structured analysis of Milk and Yogurt data, capturing spatial and temporal dependencies relevant to fermentation trends.[5] It consists of a Temporal Analysis Engine for analyzing time-series data and a Spatial Analysis Engine to examine geographical variations in fermentation prevalence.

Machine Learning Layer: *Training Module:* Uses historical data to train machine learning models like SVM, Decision Trees, and Random Forests. *Detection Module:* Applies trained models to real-time Milk data, detecting Milk and yogurt Fermentation risk factors in individuals. *Prediction Module:* Forecasts future Milk and yogurt Fermentation trends and risk patterns based on historical data and trends.

Decision & Response Layer: The Decision and response Layer assesses model predictions against quality standards and initiates quality control when subpar quality is detected. This involves real-time adjustments in both milk production and yogurt fermentation. The layer issues alerts establishes a feedback loop for continuous model improvement, and optimizes resource allocation. Data visualization aids informed decision-making, while adaptive control ensures product quality. Integration with process control systems automates adjustments, making it the bridge between data analysis and practical implementation, ensuring quality and efficiency.

Security Layer: Encryption: Implement robust data encryption protocols to secure all Milk & Yogurt data, whether in transit or at rest, ensuring the privacy of patient information.[7]**.** Authentication & Authorization: Employ security measures like multi-factor authentication and role-based access controls to ensure that only authorized personnel, including Milk and yogurt care providers and patients, have access to sensitive Milk and yogurt information within the system.

In the context of the Milk & Yogurt Fermentation detection project, these layers and security measures contribute to the correct identification of Milk & Yogurt Fermentation risk factors and personalized responses while keeping the privacy and security of Milk & Yogurt data.

VI. METRICS /MATHEMATICAL MODEL:

1. Milk & Yogurt Fermentation Detection Accuracy: In our study, we evaluated the model's performance in accurately finding individuals as obese or overweight based on their BMI and relevant indicators. The accuracy of Milk and yogurt Fermentation detection was calculated as follows:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

2. Precision and Recall for Milk & Yogurt Fermentation: Precision, which measures the percentage of true positive Milk and yogurt Fermentation predictions among all positive predictions, was computed as follows:

Precision = TP / (TP + FP)

Recall, showing the percentage of true positives among all actual obese individuals, was calculated as follows:

Recall = TP / (TP + FN)

3.F1-Score for Milk & Yogurt Fermentation: The F1-Score, which balances precision and recall for Milk and yogurt Fermentation detection, was decided using the formula:

F1-Score = 2 * (Precision * Recall) / (Precision + Recall)

4. AUC-ROC for Milk & Yogurt Fermentation: To evaluate the model's ability to distinguish between obese and non-obese individuals, we used the Area Under the Receiver Operating Characteristic (ROC) curve.

5. Body Fat Percentage Prediction Accuracy: If your model predicts body fat percentage, you can assess the accuracy of these predictions using metrics such as Mean Absolute Error (MAE) and Mean Squared Error (MSE).

MAE = Σ |Actual - Predicted| / Number of Instances

MSE = Σ (Actual - Predicted) ^2 / Number of Instances

6. Recommendation System Effectiveness: We evaluated the effectiveness of our recommendation system by measuring the percentage of personalized recommendations that individuals successfully followed.[6]

MATHEMATICAL MODEL:

Let S be the Whole system S= I, P, O

I- input, (TEXT), P- procedure, O- output

Input(I)

I = Text Dataset

Where,

Dataset- Text dataset Using SVM Algorithm

Procedure (P),

P=I, Using I System perform operations

Creating a mathematical model for Milk quality prediction and yogurt fermentation analysis using machine learning often involves the use of various algorithms and techniques.[1]

For Milk Quality Prediction:

1. Define a set of input parameters:
   - Fat Content (FC)
   - Protein Content (PC)
   - Bacterial Count (BC)
   - Temperature (T)
   - Humidity (H)
   - Source Information (SI)
   - Processing Methods (PM)

2. Develop a regression model, such as a Random Forest Regression, to predict Milk Quality (MQ) based on these inputs:

MQ = f (FC, PC, BC, T, H, SI, PM)

For Yogurt Fermentation Analysis:

1. Define a set of input parameters:
   - Milk Characteristics (MC)
   - Bacterial Strains (BS)
   - Fermentation Time (FT)
   - Temperature (T)
   - Humidity (H)
   - Chemical Properties (CP)

2. Utilize various techniques, such as Recurrent Neural Networks (RNNs) and statistical analysis, to model the fermentation process:

Yogurt Quality (YQ) = g (MC, BS, FT, T, H, CP)

VI. CONCLUSION

In conclusion, applying machine learning in the dairy industry, specifically for milk and yogurt quality prediction and fermentation analysis, can enhance efficiency, product quality, and consumer safety. Accurately distinguishing between milk and yogurt is a key challenge that can be tackled with robust machine-learning models. Success in these aims requires a multidisciplinary approach, involving domain ability, data engineering, and advanced machine learning techniques. A diverse dataset, including images, spectroscopic data, and chemical properties, is crucial for correct model training. Data preprocessing is essential for noise removal and data standardization. Various machine learning algorithms, such as support vector machines, random forests, and deep learning models like Convolutional Neural Networks (CNNs), can be used to build these models. These systems can be implemented in real-time, at production lines, dairy farms, or retail settings, ensuring prompt action in cases of mislabelling or contamination, and keeping product authenticity and quality.

REFERENCE

[1] S Louis and K Don," Milk production estimates using feed-forward artificial neural networks", Computers and Electronics in Agriculture, vol. 32, no. 1, pp. 21-30, 2023.

[2] Doran, P. Bioprocess Engineering Principles, 2nd ed.; Academic Press: London, UK, 2021. [Google Scholar]

[3] S. Stankowski, G. Ostojic, I. Senk, M. Rakic-Skokovic, S. Trivunovic, and D. Kucevic, “Dairy cow monitoring by RFID,” Scientia Agricola, vol. 69, no. 1, pp. 75–80, 2020

[4] Wang, Q.; Parsons, R.; Zhang, G. China’s dairy markets: Trends, disparities, and implications for trade. China Agri. Econ. Rev. 2010, 2, 356–371. [Google Scholar] [CrossRef].

[5] Blanc, B., and Odet, G. 1981. Appearance flavor and texture aspects: Recent development. Azevedo et al., 2023 C.F. Azevedo, M.D.V. de Resende, F.F. E Silva, J.M.S.

[6] Viana, M.S.F. Valente, M.F.R. Resende Jr., P. Munoz Ridge, LASSO and Bayesian additive-dominance genomic models BMC Genet., 16 (2023), p. 105 https://doi.org/10.1186/s12863-015-0264-2.

[7] M. Ghobakhloo, The future of manufacturing industry: A strategic roadmap toward industry 4.0, J. Manuf. Technol. Manag., vol.29, no.6, pp.910-936, 2021.