International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

A Neural Network Framework for Forecasting Milk Fat and SNF from Cow Feeding and Environmental Parameters

B. Veerabahu Pillai
Research Scholar
Dept. of Computer Science & Engg.
Hindustan Institute of Tech. & Science
Chennai, Tamil Nadu, India
Orcid ID : 0009-0001-9714-2173

Dr. T. Sudalaimuthu
Professor (Supervisor / Guide)
Dept. of Computer Science & Engg.
Hindustan Institute of Tech. & Science
Chennai, Tamil Nadu, India
Orcid ID : 0000-0003-0371-9371

Dr. N. Muthuvairavan Pillai
Assoc. Prof. (Research Supervisor)
Dept. of Computer Science & Engg.
R.M.D. Engineering College
Chennai, Tamil Nadu, India

Abstract: Milk quality is a critical factor in the dairy industry because it directly influences nutritional value, processing characteristics, economic returns, and consumer acceptance. Among the different components used to assess milk quality, milk fat and solids-not-fat (SNF) are particularly important because their concentrations reflect the nutritional status of the animal and contribute significantly to the physical and commercial properties of milk. However, these components are affected by a combination of factors rather than a single variable. Feed composition, quantity of feed consumed, nutrient availability, lactation stage, breed, age, body condition, milk yield, ambient temperature, relative humidity, and environmental stress can interact with each other and produce complex changes in milk composition. Conventional estimation of milk fat and SNF commonly relies on laboratory-based testing, periodic sampling, or simplified statistical relationships. Although these methods can provide reliable measurements, they may be time-consuming, resource-intensive, and less suitable for continuous or early-stage prediction in modern dairy management systems.To address these limitations, this study proposes a neural network-based framework for forecasting milk fat and SNF from cow feeding and environmental parameters. The proposed framework considers feeding-related variables and environmental conditions as predictive inputs and establishes a nonlinear relationship between these factors and the resulting milk composition. The data processing pipeline includes data collection, quality checking, handling of missing or inconsistent observations, feature preparation, normalization, and separation of the dataset into appropriate training and testing subsets. A neural network model is subsequently trained to identify complex patterns within the input variables and generate predictions for milk fat and SNF. The framework is designed to support multi-output prediction, allowing both milk-quality parameters to be estimated from the same set of feeding and environmental observations.The performance of the proposed model can be assessed using widely accepted regression metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). In addition to prediction performance, analysis of the input variables can provide useful insights into the factors associated with variations in milk composition. The proposed framework provides a foundation for moving from reactive milk-quality assessment toward predictive dairy management, where feeding and environmental conditions can be evaluated before their potential effects on milk quality become evident. Such a system can assist dairy farmers, nutritionists, and researchers in improving feeding strategies, identifying unfavorable environmental conditions, and supporting more consistent milk production. Furthermore, the framework can be extended to incorporate real-time sensor measurements from smart dairy environments, enabling continuous data acquisition and near-real-time forecasting. Overall, the proposed neural network approach demonstrates the potential of artificial intelligence and data-driven modeling to support precision dairy farming and improve the efficiency, consistency, and sustainability of milk-quality management.

Keywords : Milk Fat; Solids-Not-Fat (SNF); Neural Networks; Milk Quality Prediction; Machine Learning; Artificial Intelligence; Cow Feeding; Feed Intake; Environmental Parameters; Temperature-Humidity Index; Dairy Cattle; Precision Dairy Farming; Predictive Modeling; Animal Nutrition; Smart Dairy Management.

I. INTRODUCTION

The dairy industry is a major contributor to the global food system, supplying milk and numerous products that provide essential nutritional value. The quality of milk, however, is not determined by a single factor. It is the result of interactions among the animal’s biological characteristics, nutritional status, management practices, and environmental conditions.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

Within the overall composition of milk, milk fat and solids-not-fat (SNF) are two important parameters that receive considerable attention because of their nutritional contribution, influence on processing characteristics, economic significance, and role in determining milk quality. For dairy producers and processing industries, maintaining these components within desirable ranges is therefore an important aspect of effective milk production. In practice, however, their concentrations can fluctuate considerably depending on factors such as feed management, physiological status, lactation stage, and environmental conditions.Milk fat is particularly sensitive to changes in the nutritional and metabolic condition of dairy cattle. The composition of the diet and the quantity of feed consumed can affect rumen fermentation and the availability of nutrients required for milk synthesis. Dietary characteristics such as the proportion of forage and concentrate, dry matter intake, available energy, protein supply, and feeding patterns can consequently influence the amount of fat produced in milk. SNF represents another important group of milk constituents and consists primarily of proteins, lactose, minerals, and other non-fat components. Its concentration may be influenced by nutrient availability, metabolic processes, milk yield, and the physiological condition of the animal. Because both fat and SNF are influenced by several factors operating simultaneously, accurate forecasting requires a combination of relevant variables instead of depending on an individual feeding parameter.In addition to nutrition, the surrounding environment can substantially affect dairy cattle performance and milk production. Conditions characterized by high temperature and humidity can expose animals to heat stress, potentially affecting their feed intake, water consumption, metabolism, and productive efficiency. The Temperature-Humidity Index (THI) is commonly used to represent the combined thermal burden imposed by temperature and humidity and can therefore serve as a useful indicator of heat-load conditions. When environmental conditions become unfavorable, cattle may alter their feeding patterns, physiological responses, and general behavior in an effort to regulate body temperature. These responses can indirectly affect both milk yield and its composition. Consequently, environmental information should be analyzed alongside nutritional and animal-related variables when attempting to forecast changes in milk quality.Conventional determination of milk fat and SNF generally depends on laboratory procedures or dedicated milk-analysis equipment. Although these methods can provide reliable measurements, they typically assess milk composition after production and sampling have already occurred. Conducting frequent measurements may additionally involve specialized equipment, operational costs, labor, and processing time. From the viewpoint of farm management, a predictive method capable of estimating milk composition using information that is routinely collected could provide a useful supplementary decision-support tool. Forecasting changes in advance may enable farmers and livestock managers to consider appropriate modifications to feeding strategies or environmental management before significant changes in milk composition become evident.The rapid development of artificial intelligence (AI) and machine learning has expanded the possibilities for analyzing agricultural and livestock-production data. Conventional statistical techniques often rely on predefined assumptions about the relationships between input and output variables. Neural networks, in contrast, are capable of learning complex and nonlinear relationships directly from observed data. This characteristic is particularly relevant to dairy production because the influence of nutrition, environmental conditions, animal characteristics, and milk composition is rarely governed by a simple linear relationship. By identifying patterns within historical observations, a neural-network model can learn a predictive mapping between the conditions experienced by dairy cattle and the resulting characteristics of their milk.Based on this premise, the present study introduces a neural network framework for forecasting milk fat and SNF using cow feeding and environmental parameters. The proposed framework brings together relevant nutritional and environmental observations within a unified, data-driven prediction pipeline. Depending on data availability, feeding-related inputs may include feed intake, fodder quantity, dietary composition, nutrient-related measurements, and other relevant feeding attributes. Environmental inputs can include variables such as ambient temperature and relative humidity, together with calculated indicators such as THI. Information describing individual cattle may also be incorporated where available, allowing the model to account for differences between animals. Before model development, the collected data undergo appropriate preprocessing operations, including data-quality checking, treatment of missing observations, normalization or scaling, and preparation of relevant input features.The primary objective of the framework is to forecast two important indicators of milk composition, namely milk fat and SNF. A multi-output neural-network architecture can be adopted so that a common set of feeding, environmental, and animal-related inputs can be used to estimate both target variables. Such an approach allows the model to learn potentially shared relationships between the input conditions and the two milk-composition characteristics. Model effectiveness can subsequently be assessed using commonly applied regression measures, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). Together, these metrics provide an assessment of prediction error as well as the extent to which the

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

model captures the observed variation in milk fat and SNF.The practical value of the proposed framework is not limited to achieving low prediction errors. An effective forecasting system could assist dairy farmers and livestock managers in making more informed management decisions by providing an indication of how current feeding and environmental conditions may be associated with future milk-quality changes. Such information could be particularly valuable when integrated with sensor-based dairy monitoring systems. For example, environmental sensors can continuously record temperature and humidity, while feeding and production records can provide complementary information. Combining these data sources with an AI-based forecasting model could support continuous estimation of milk-composition trends and potentially generate early warnings when conditions associated with undesirable changes are identified.The proposed work also aligns with the growing field of precision livestock farming, where continuous data collection and computational analysis are increasingly used to improve animal-management practices. Rather than depending exclusively on periodic measurements or manual observation, precision livestock systems aim to support management decisions through timely and objective information. Integrating nutritional and environmental parameters within a common predictive framework provides an opportunity to develop a more comprehensive representation of the factors affecting milk composition. In this regard, the proposed neural-network approach is intended as a practical and scalable foundation for forecasting milk fat and SNF from routinely obtainable dairy-farm data.Future development of the framework could involve the use of larger and more diverse datasets collected across different farms, breeds, seasons, and management conditions. Additional animal-level characteristics could further improve the model's ability to represent individual variation. The framework could also be connected with Internet of Things (IoT) sensors to facilitate continuous acquisition of environmental and production-related information. With sufficient data and validation, more advanced deep-learning architectures could be investigated to improve forecasting performance and robustness. Ultimately, the integration of AI, IoT, nutritional monitoring, and environmental sensing could contribute toward intelligent dairy-management systems capable of providing timely information for maintaining milk quality and improving production efficiency.

II. BACKGROUND AND RELATED WORK

A.Traditional Milk-Quality Prediction

Milk-quality evaluation has historically been based on the direct analysis of physicochemical properties obtained from representative milk samples. Commonly examined parameters include milk fat, protein, lactose, solids-not-fat (SNF), density, acidity, and total solids. These characteristics are usually measured through established laboratory techniques or dedicated milk-analysis instruments. Such measurements remain important in the dairy industry because they provide objective information about milk composition and are routinely used for quality assurance, pricing, processing decisions, and regulatory compliance.Despite their reliability, conventional measurement techniques have a primarily retrospective nature. The composition of milk is determined only after the milk has been produced and a sample has been collected. Consequently, these methods provide limited information about the conditions that may have contributed to a subsequent change in milk quality. Repeated sampling can also involve additional requirements related to sample handling, laboratory facilities, trained personnel, testing time, and operational cost. From the perspective of dairy management, therefore, obtaining a measurement of milk quality is useful, but being able to anticipate potential changes before they become evident could provide greater practical value.Statistical techniques have traditionally been employed to examine the relationship between milk composition and different animal- and management-related variables. Factors such as feed consumption, breed, age, lactation stage, milk yield, and nutritional status have been considered when studying variations in milk components. Regression models, for example, can be used to quantify the association between selected explanatory variables and specific milk-quality characteristics. Although these methods offer useful interpretations and can establish important relationships, they commonly depend on assumptions concerning the mathematical form of the relationship between the input variables and the target parameter. In dairy production, however, milk synthesis is governed by several biological and environmental processes that may interact in nonlinear ways. This has created increasing interest in more flexible, data-driven approaches.

B. Machine-Learning Approaches

The growth of digital technologies in agriculture has resulted in the generation of increasingly large datasets from farm records, automated equipment, and environmental sensors. These developments have created new possibilities for applying machine-learning techniques to dairy production and livestock management. Unlike conventional analytical methods that generally examine predefined relationships, machine-learning algorithms can process several variables simultaneously and identify patterns from previously observed data. Methods including linear regression, decision trees, random forests, support vector regression, k-nearest-neighbor algorithms, and ensemble techniques have consequently been explored for different livestock and milk-production

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

prediction tasks.One of the main advantages of machine learning is its ability to represent relationships that may not follow a simple linear pattern. For instance, an increase in feed intake does not necessarily result in a proportional increase in milk-fat concentration. The response can vary according to dietary composition, animal condition, lactation stage, and nutrient utilization. Environmental factors exhibit similar complexity. The effect of increased temperature, for example, may become more pronounced when relative humidity is also high or when the animal has limited ability to dissipate excess body heat. Machine-learning models can consider such variables together, providing a more flexible representation of the conditions associated with milk-quality variation.However, the effectiveness of a machine-learning model is closely related to the characteristics of the dataset used for its development. Incomplete records, measurement inconsistencies, seasonal variation, differences among individual animals, and variations in farm-management practices can all affect the ability of a model to perform reliably on unseen data. Appropriate preprocessing, feature selection, normalization, and data-quality management are therefore important stages in developing a robust prediction system. In addition, models that depend only on production-related measurements may fail to account for nutritional and environmental conditions that contribute to changes in milk composition.

C. Neural-Network Approaches

Neural networks have emerged as an important class of machine-learning models for applications where the relationship between input variables and predicted outcomes is complex. A typical neural network contains an input layer, one or more hidden layers, and an output layer consisting of interconnected computational units. During training, the model modifies its internal weights and biases based on the difference between its predictions and the observed target values. Through repeated optimization, the network can learn patterns from multidimensional data without requiring an explicit mathematical equation describing the relationship between every input and output.The use of neural networks has expanded across several agricultural and livestock applications, including milk-yield estimation, animal-health monitoring, disease identification, behavior analysis, feed-management applications, and production forecasting. Their ability to process multiple variables makes them particularly relevant to dairy systems, where nutritional, physiological, management, and environmental factors can influence one another. Neural-network architectures can also be configured for multi-output prediction, enabling multiple related milk characteristics to be estimated from a common set of input variables.Nevertheless, neural networks are not automatically effective simply because they are capable of representing complex relationships. Their predictive performance can be affected by the amount and quality of training data, network architecture, feature scaling, optimization strategy, and hyperparameter configuration. An excessively complex model may memorize variations that are specific to the training dataset, while a model with insufficient capacity may fail to capture meaningful patterns. Appropriate training, validation, regularization, and independent testing are therefore necessary to assess whether the learned relationships can generalize beyond the observations used during model development.In the context of milk-quality forecasting, neural networks are particularly promising because the composition of milk reflects the conditions under which the animal is managed and produces milk. By providing information related to feeding, animal characteristics, and environmental conditions, a neural network can potentially learn combinations of factors associated with changes in milk fat and SNF. This makes neural-network modeling a suitable foundation for developing predictive tools that can complement conventional dairy-quality assessment.

D. Feed-Related Milk Composition Prediction

Feed management has a direct relationship with dairy-cattle productivity because the nutrients supplied through the diet provide the resources required for maintenance, reproduction, growth, and milk synthesis. Consequently, both the quantity of feed consumed and its nutritional composition are important when examining variations in milk composition. Parameters such as dry matter intake, forage proportion, concentrate consumption, dietary energy, crude protein, fiber content, and feeding frequency can provide useful information for understanding differences in milk-production characteristics.Milk fat is particularly influenced by nutritional conditions because processes occurring within the rumen generate metabolites that contribute to milk-fat synthesis. Alterations in the balance between forage and concentrate, dietary fiber availability, and overall nutrient composition can modify rumen fermentation and consequently affect the concentration of fat in milk. SNF can also respond to nutritional and physiological conditions because it represents a group of milk constituents that includes protein, lactose, minerals, and other non-fat solids.Previous predictive studies have examined the relationship between nutritional variables and milk yield or individual milk components. However, predicting milk composition from feed information alone remains challenging because cattle do not necessarily respond identically to the same diet. Differences in breed, age, parity, lactation stage, body condition, health, and individual nutrient utilization can lead to variations in the resulting milk composition. Incorporating animal-related and environmental information together with feeding

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

variables can therefore provide a broader representation of the factors affecting milk-quality characteristics.The increasing availability of structured feeding records also creates an opportunity to move beyond retrospective nutritional analysis. Instead of examining whether a particular dietary component is associated with milk fat or SNF in isolation, machine-learning methods can learn relationships among several nutritional variables simultaneously. Such predictive capability is relevant to precision dairy farming, where management decisions can increasingly be informed by data collected at the individual-animal level.

E. Environmental-Stress Effects

The surrounding environment can significantly affect dairy cattle because environmental conditions influence behavior, feed consumption, water requirements, physiological regulation, and productive performance. Among the different environmental variables, ambient temperature and relative humidity are particularly important because their combined effect determines the level of thermal stress experienced by an animal. When heat load becomes excessive, cattle may change their activity and feeding behavior as part of their effort to maintain an appropriate body temperature.Heat stress is of particular concern in regions characterized by high temperatures and humidity. Under such conditions, changes in respiration, water consumption, activity, feed intake, and metabolic processes can reduce production efficiency. A reduction in feed intake is especially relevant to milk composition because fewer nutrients may subsequently be available for milk synthesis. The Temperature-Humidity Index (THI) is therefore commonly used as an indicator of combined temperature and humidity conditions and can serve as an informative input when developing predictive dairy models.Environmental stress should not be considered independently from nutrition. A change in temperature may influence feed intake, while the resulting reduction in nutrient availability can affect milk synthesis. Similarly, the effect of environmental heat may vary according to the animal's nutritional status, production level, physiological condition, and ability to tolerate thermal stress. Examining temperature or humidity individually may consequently provide an incomplete representation of the conditions affecting milk quality. Combining environmental variables with feeding information allows a prediction model to investigate their potential interactions.Modern dairy farms increasingly employ sensors and automated monitoring technologies capable of recording environmental conditions at regular intervals. Temperature, humidity, and related measurements can therefore be collected alongside feeding records and milk-quality observations. Integrating these sources of information creates an opportunity to develop predictive models that identify relationships between environmental conditions and subsequent changes in milk composition. This integration represents an important step toward connecting conventional dairy monitoring with intelligent precision-livestock systems.

F. Research Gap

Existing research indicates that milk composition is influenced by a broad combination of nutritional, physiological, environmental, and management-related factors. Conventional laboratory analysis continues to provide reliable measurements of milk constituents, but its primary purpose is to determine the composition of milk that has already been produced. Statistical models can help identify relationships between milk characteristics and selected explanatory variables, while machine-learning techniques provide greater flexibility when the relationships become nonlinear. Neural networks extend this capability by learning complex patterns involving several variables simultaneously. Nevertheless, important research opportunities remain. A considerable portion of existing predictive work concentrates on milk yield or individual production indicators rather than jointly estimating milk-fat and SNF levels. In addition, nutritional factors and environmental conditions are not always incorporated within the same predictive framework, despite the possibility that heat stress and feeding behavior can influence one another. Another concern is model generalization. A model developed using data from a restricted set of animals, seasons, feeding regimes, or environmental conditions may not perform equally well when applied to different production environments. There is therefore a need for forecasting approaches that can combine routinely available farm-level information while accounting for the complex conditions under which milk is produced.To address these limitations, the present study develops a neural-network-based framework for forecasting milk fat and SNF using cow feeding and environmental parameters. The proposed approach considers milk composition as an outcome influenced by interacting nutritional and environmental conditions rather than as an isolated production variable. Relevant input data are first prepared through appropriate preprocessing procedures and are then provided to a neural network capable of learning nonlinear relationships within the dataset. The resulting model is used to estimate both milk-fat and SNF levels from the available input conditions.The proposed framework also provides a basis for future integration with sensor-enabled dairy-management systems. Feeding records, environmental measurements, and other farm observations could potentially be collected continuously and supplied to predictive models, allowing milk-quality trends to be estimated without depending exclusively on periodic laboratory

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

assessment. Such a development could support earlier management intervention and improve the use of data in dairy production.The principal contribution of this work is the integration of feeding-related and environmental variables into a unified neural-network framework for the joint forecasting of milk fat and SNF. This approach shifts the emphasis from measuring milk composition only after production toward estimating potential changes from conditions that can be observed during the production process. By combining nutritional and environmental information within a single predictive framework, the study provides a foundation for more proactive dairy-quality monitoring and contributes to the broader development of data-driven precision livestock farming.

III. RESEARCH GAP AND PROBLEM FORMULATION

Existing dairy prediction approaches have mainly focused on individual aspects such as milk yield, animal health, feeding management, or milk-quality estimation, while the combined influence of nutritional, animal-related, environmental, and production factors is not always considered within a single predictive framework. Conventional statistical and machine-learning methods may also have limitations when the relationship between these variables and milk composition is nonlinear and involves interactions among multiple factors. In addition, many existing approaches predict individual milk-quality parameters separately, which may not fully utilize the shared information between related milk components. To address these limitations, this study proposes a Multi-Output Multi-Layer Perceptron Deep Neural Network (MLP-DNN) for simultaneous prediction of milk fat and solids-not-fat (SNF).

The input feature vector for a dairy cow at time t is represented as:

$$\mathbf{X_t = [x_{t1}, x_{t2}, x_{t3}, \ldots, x_{tn}]^T}$$

where X_t represents the input feature vector at time t, x_{tj} represents the j-th input feature, and n denotes the total number of input variables. The selected features may include nutritional parameters such as feed intake and fodder quantity, animal-related characteristics, environmental parameters such as ambient temperature and relative humidity, derived parameters such as Temperature-Humidity Index (THI), and milk-production information such as milk yield. These variables provide the neural network with the information required to learn the complex relationships between dairy-farm conditions and milk-quality characteristics.

The proposed neural-network model can be mathematically represented as:

$$\mathbf{\hat{Y}_t = f\theta(X_t)}$$

where fθ represents the proposed MLP-DNN with learnable parameters θ, X_t represents the input feature vector at time t, and $\hat{Y}_t$ represents the predicted output vector. Since the proposed study is formulated as a multi-output regression problem, the output vector consists of two milk-quality parameters and can be expressed as:

$$\mathbf{\hat{Y}_t = [\ \hat{Y}Fat_{,t}\ ;\ \hat{Y}SNF_{,t}\ ]}$$

where $\hat{Y}Fat_{,t}$ represents the predicted milk fat percentage and $\hat{Y}SNF_{,t}$ represents the predicted SNF percentage. The model therefore learns a shared nonlinear representation from the input variables and produces both milk-quality predictions simultaneously.For N observations, the prediction error is minimized using the Mean Squared Error (MSE) loss function, which is defined as:

$$\mathbf{L_{mse} = 1/(2N)\ \Sigma_{i=1}^{N}\ [(y_i,Fat - \hat{y}_i,Fat)^2 + (y_i,SNF - \hat{y}_i,SNF)^2]}$$

where N represents the total number of observations, y_i,Fat and y_i,SNF are the actual milk fat and SNF values for observation i, and $\hat{y}_i$,Fat and $\hat{y}_i$,SNF are the corresponding predicted values. The factor 1/2 is included to simplify the gradient calculation during optimization. Minimizing this loss enables the neural network to adjust its weights and biases so that the predicted milk-quality values become closer to the measured values.The problem can therefore be formulated as finding the optimal model parameters θ* that minimize the prediction loss over the training data:

$$\theta = \arg\min\theta\ L_{mse}(Y, f\theta(X))*$$

where θ* represents the optimal set of trainable neural-network parameters, X represents the input dataset, and Y represents the corresponding actual milk-quality observations. The optimization process aims to minimize the prediction error while maintaining good generalization performance on unseen observations.

Therefore, the research problem addressed in this study is to develop a nonlinear multi-output learning model that can effectively utilize nutritional, animal-related, environmental, and production information to simultaneously predict milk fat and SNF with low prediction error and reliable predictive performance. Unlike approaches that independently estimate individual milk-quality parameters, the proposed

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

MLP-DNN learns a shared representation of the input features and generates both outputs from the same network, providing an integrated framework for dairy milk-quality prediction.

IV. PROPOSED METHODOLOGY

A. *Overall Architecture*

The proposed framework is designed to forecast milk fat and solids-not-fat (SNF) by integrating information related to cow feeding, animal characteristics, and environmental conditions. The overall architecture establishes a systematic pathway through which the collected data are transformed into meaningful predictions using a neural-network model.The framework begins with the collection of relevant feeding, cow, environmental, and milk-production data. Feeding-related information may include feed quantity, feed composition, concentrate intake, fodder intake, protein intake, and energy-related variables. Cow-specific information such as age, breed, body weight, parity, and lactation stage can also be incorporated to account for differences among individual animals. Environmental conditions, particularly temperature and humidity, are considered because changes in the surrounding environment can influence feed intake, animal comfort, metabolic activity, and ultimately milk composition.The collected parameters are combined into a structured dataset and subjected to data preprocessing. This stage includes handling missing observations, identifying abnormal values, maintaining data consistency, and scaling numerical variables to suitable ranges. The processed dataset is then passed through a feature-engineering stage, where relevant variables are selected or transformed to improve the ability of the model to identify relationships between the input conditions and milk composition.The resulting feature set is provided to the proposed neural-network model. The network learns nonlinear relationships between cow feeding patterns, animal characteristics, environmental conditions, and the resulting milk-quality characteristics.

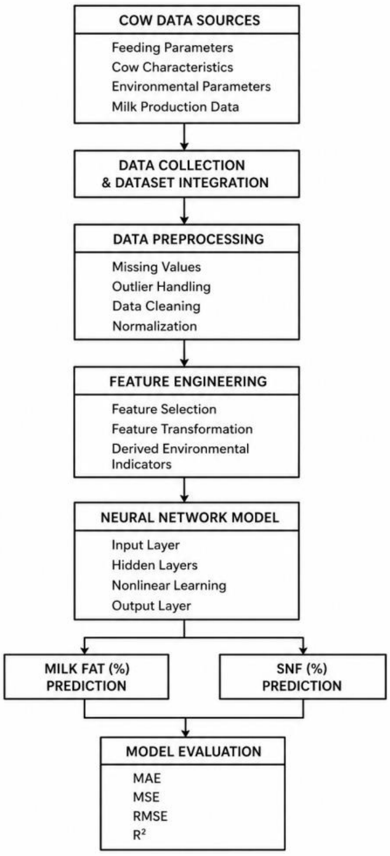

Figure 1: Overall Block Diagram of the Proposed Framework

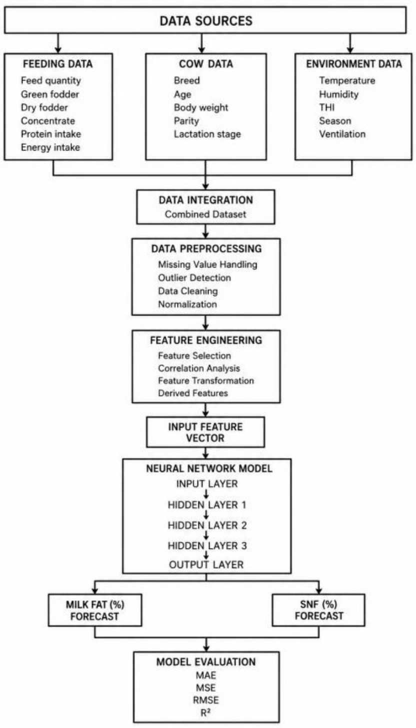

Figure 2: Detailed System Architecture of the Proposed Framework

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

B. Data Collection

The quality of a predictive model is closely related to the quality and representativeness of the dataset used for its development. In the proposed framework, data collection is designed to capture the major factors that may contribute to variations in milk composition, with particular emphasis on milk fat and solids-not-fat (SNF). Since milk composition is influenced by more than one aspect of dairy production, the dataset combines information from feeding management, individual animal characteristics, environmental conditions, milk production, and milk-quality measurements.The data-collection process is organized so that the different measurements correspond to the same cow and observation period. This synchronization is important because feeding conditions, environmental exposure, animal status, and milk composition can change over time. Associating these variables correctly allows the neural network to learn meaningful relationships rather than patterns created by mismatched observations. The collected data are subsequently consolidated into a structured dataset for preprocessing, feature engineering, and model development.

C. Sources of Cow Feeding Data

Feeding data are collected to represent the nutritional conditions experienced by individual dairy cows. Depending on the availability and accuracy of farm records, the dataset can contain information regarding the quantity and type of feed supplied to each animal. Important variables may include green fodder quantity, dry fodder quantity, concentrate feed quantity, total feed intake, feeding frequency, dry matter intake, crude protein, dietary energy, and fiber-related parameters.The quantity of each feed component provides information about the overall nutritional supply available to the animal, whereas nutrient-related variables provide additional information about the quality of the diet. For example, two cows may receive similar quantities of feed but differ in nutrient intake because of differences in feed composition. Including both quantity-related and nutritional information can therefore provide a more representative description of feeding conditions.Where direct measurements of nutrient intake are unavailable, estimated values can be obtained from available feed-composition records and feeding quantities, provided that the estimation procedure is consistently applied. Such estimates should be clearly distinguished from directly measured values to avoid introducing unnecessary uncertainty into the dataset.Feeding records are linked to individual cows using a unique animal identifier. The date and observation period are also retained so that feeding information can be associated with the corresponding milk-production and quality measurements. This temporal association enables the model to learn how feeding conditions during a particular period relate to the milk produced during the corresponding period.Repeated feeding observations are preferable to isolated measurements because they allow changes in dietary conditions to be represented over time. This is particularly relevant for dairy cattle because feed composition and intake may change according to lactation stage, seasonal availability, management practices, or changes in production requirements.

D. Collection of Environmental Parameters

Environmental information is collected to represent the external conditions experienced by the animals during milk production. The primary environmental parameters considered in the proposed framework are ambient temperature and relative humidity. These measurements can be obtained using environmental sensors positioned within or around the cattle housing area. Temperature measurements provide an indication of the thermal conditions surrounding the animals, while relative humidity describes the amount of moisture present in the surrounding air. Considering these variables together is important because high temperature combined with high humidity can create a greater thermal burden than either condition alone. The collected temperature and humidity values can be used to calculate the Temperature-Humidity Index (THI). THI provides a combined representation of thermal conditions and can be used as an additional model input. Depending on the available data, the environmental dataset may also include information such as season, time of measurement, housing type, ventilation condition, and other relevant environmental observations. Environmental data should preferably be collected at a sufficiently high frequency to capture short-term variations. Where measurements are recorded at intervals different from the feeding or milk-quality records, an aggregation procedure can be applied to obtain representative values for the corresponding observation period. For example, average, minimum, maximum, or accumulated environmental measurements may be calculated depending on the intended modeling strategy .Environmental measurements are subsequently synchronized with the corresponding cow and production records. This ensures that the model receives environmental information that realistically represents the conditions experienced during the period associated with each milk-quality observation.

E. Collection of Cow-Related Information

Animal-specific information is incorporated to represent biological differences among individual cows. The dataset may include variables such as breed, age, body weight, parity, lactation number,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

lactation stage, days in milk, and other relevant production characteristics. These attributes provide additional context for interpreting differences in milk composition between animals.The inclusion of animal-related information is important because dairy cows do not respond identically to the same feeding or environmental conditions. Differences in physiological status, breed characteristics, production level, and lactation stage may influence nutrient utilization and milk synthesis. Therefore, an environmental or feeding condition that produces a particular response in one animal may not necessarily result in the same response in another. Each cow is assigned a unique identification value. This identifier is used to connect repeated observations collected from the same animal over different dates or production periods. Maintaining this structure makes it possible to distinguish between differences occurring across individual cows and changes occurring within the same cow over time. Where repeated measurements are available, the dataset can represent the progression of individual animals through different stages of lactation. This provides additional information for the neural network and can improve its ability to distinguish physiological changes from variations caused by feeding or environmental conditions.

F. Collection of Milk Production and Quality Measurements

Milk-production and milk-quality measurements form the target component of the dataset. Milk yield can be recorded on a daily basis or at predefined sampling intervals depending on the available farm-management system. The production measurement is retained together with the milk-quality values corresponding to the same cow and observation period. The primary target variables of the proposed framework are milk fat percentage (%) and SNF percentage (%). These parameters are selected because they represent important characteristics of milk composition and are influenced by both nutritional and physiological conditions. Milk-quality measurements should preferably be obtained using standardized laboratory methods or properly calibrated milk-testing equipment. Consistent measurement procedures are important because systematic differences in measurement technique can introduce artificial variation into the target values. Each milk-quality record is associated with the corresponding cow ID, measurement date, milk yield, fat percentage, and SNF percentage. Additional quality indicators such as protein or lactose may also be retained when available, even if they are not used as prediction targets. Such information can be useful for subsequent analysis or future extensions of the framework. The timing of milk sampling is also considered during dataset preparation. Measurements collected at substantially different time periods from the associated feeding or environmental observations may not accurately represent the same production condition. Therefore, observations are matched according to a defined temporal window before being included in the final modeling dataset.

G. Input Parameters

The proposed neural-network framework uses a combination of feeding, cow-related, environmental, and production parameters to forecast milk fat and SNF. The selection of input variables is based on their potential influence on milk production and composition. The input data are organized into meaningful categories to enable the model to capture both nutritional and environmental effects on milk quality.

H. Feeding Parameters

Feeding parameters represent the nutritional inputs provided to the cow and describe both the quantity and composition of feed consumed during the observation period. In the proposed framework, these parameters include the quantity of green fodder, dry fodder, and concentrate feed, along with the total feed intake and feeding frequency. Nutritional characteristics such as crude protein intake, energy intake, and dietary composition are also considered whenever reliable information is available. In addition, dry matter intake (DMI) is incorporated to represent the amount of feed consumed after accounting for its moisture content. Collectively, these feeding-related variables provide important information about the nutritional conditions of the animal and enable the neural-network model to learn their potential relationships with milk fat and SNF composition.

I. Cow-Related Parameters

Cow-related parameters are incorporated into the proposed framework to account for biological and physiological differences among individual animals. These characteristics provide additional context for understanding variations in milk production and composition across cows. The considered parameters include **breed, age, body weight, parity, lactation stage, and days in milk**. Breed can reflect inherent differences in production characteristics, while age and body weight provide information about the animal’s physical condition. Parity represents the number of previous calvings and can be associated with changes in productive performance, whereas lactation stage and days in milk help capture variations in milk composition throughout the lactation cycle. Together, these cow-specific parameters enable the neural-network model to account for individual animal characteristics when forecasting milk fat and SNF.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

J. Environmental Parameters

Environmental conditions are incorporated into the proposed framework because variations in the surrounding environment can influence animal comfort, feed intake, physiological responses, and milk production. The primary environmental parameters considered in the study include ambient temperature (°C), relative humidity (%), Temperature-Humidity Index (THI), season, and ventilation conditions, where such information is available. Ambient temperature and relative humidity provide direct measures of the environmental conditions experienced by the animals, while THI combines these factors to represent the potential effect of heat and humidity on animal comfort. Seasonal information is included to capture broader environmental variations throughout the year, and ventilation conditions can provide additional information about the effectiveness of air movement within the housing environment. These parameters enable the neural-network model to account for environmental variations when learning the relationship between external conditions and changes in milk fat and SNF.

K. Milk Production Parameters

Milk production variables provide additional information about the production state of the cow. The principal production parameter considered as an input is: Daily milk yield (litres/day)Including milk yield alongside feeding and environmental variables enables the neural network to learn relationships between production level and milk composition.

L. Target Parameters

The proposed model performs a multi-output prediction, with two milk-quality characteristics serving as the target variables: Milk Fat (%) ,Solids-Not-Fat (SNF) (%).The model receives the selected feeding, cow-related, environmental, and production parameters as inputs and simultaneously estimates milk fat and SNF.

Category	Parameters	Role
Feeding	Green fodder, dry fodder, concentrate, total feed intake, DMI	Input
Nutrition	Protein intake, energy intake	Input
Cow-related	Breed, age, body weight, parity, lactation stage, days in milk	Input
Environmental	Temperature, humidity, THI, season, ventilation	Input
Production	Daily milk yield	Input
Milk quality	Milk fat (%)	Target
Milk quality	SNF (%)	Target

Table 1. Input and Output Parameters Used in the Proposed Framework

M. Data Preprocessing

Data preprocessing is carried out to improve the quality and consistency of the collected data before using it to train the neural-network model. The collected dataset may contain missing values, duplicate records, incorrect measurements, and different units. These problems are checked and corrected during preprocessing to reduce errors and improve the reliability of the prediction model First, the dataset is examined to identify missing values and duplicate observations. Missing values in important input variables are handled using suitable methods such as mean, median, or interpolation, depending on the type of data. If the target values, such as milk fat or SNF, are missing and cannot be reliably recovered, those records are removed from the training dataset. Duplicate records are also removed to prevent the same observation from affecting the model more than once.The collected values are then checked for incorrect or unusual measurements. Outliers are identified by examining the range and distribution of the variables. However, not every unusual value is removed because differences in feed intake, environmental conditions, milk yield, and milk composition can occur naturally between cows. Only values that are clearly caused by measurement or recording errors are removed or corrected.The different measurement units used in the dataset are converted into a consistent format. For example, feed quantity, milk yield, temperature, and other numerical values are represented using common units. Categorical information such as breed and season is converted into numerical form using suitable encoding methods so that it can be processed by the neural network. Since the input variables have different numerical ranges, feature scaling is applied before model training. Variables such as feed quantity, temperature, humidity, milk yield, and other measurements may have different scales.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

Standardization or normalization is therefore used to bring the numerical features into a suitable range. Standardization can be calculated using:

$$z_i = \frac{x_i - \mu}{\sigma}$$

where x_i is the original value, μ is the mean of the feature, and σ is its standard deviation. This scaling helps the neural network learn more efficiently and prevents variables with larger numerical values from having an unnecessary influence on the training process. The data collected from different sources are also matched using the cow identification number and observation date or period. This ensures that feeding information, environmental conditions, animal characteristics, milk yield, milk fat, and SNF belong to the correct cow and corresponding observation period. Environmental measurements such as temperature and humidity may be averaged over the relevant observation period when they are recorded more frequently than milk-quality measurements. After cleaning and preparing the data, the dataset is divided into training, validation, and testing sets. The training set is used to train the neural network, the validation set is used to monitor and improve the model during development, and the testing set is used for the final evaluation of prediction performance. Care is taken to avoid using information from the testing data during training or preprocessing.The final preprocessed dataset contains the selected feeding, cow-related, and environmental parameters as input variables, while milk fat (%) and SNF (%) are used as the main output variables.

N. Feature Engineering

Feature engineering is performed to prepare the processed data in a useful form for the neural-network model. The available feeding, cow-related, environmental, and milk-production variables are examined to determine which features are relevant for predicting milk fat and SNF. Features that provide useful information are retained, while variables that are repeated, unnecessary, or have very little useful information may be removed.New features can also be created from the existing variables when they provide additional information to the model. For example, Temperature-Humidity Index (THI) can be calculated using ambient temperature and relative humidity to represent the combined effect of environmental heat conditions. Other useful features may include total feed intake, combined fodder quantity, nutrient-related values, or other calculated measurements when sufficient data are available.The relationship between the input variables and the target values is also examined to understand which features may be associated with changes in milk fat and SNF. Correlation analysis can be used to identify the strength and direction of relationships between numerical variables. Features that provide very similar information may be reduced to avoid unnecessary repetition. Feature-importance methods can also be used to identify variables that contribute more strongly to the prediction results.Interactions between feeding and environmental conditions may also be considered because their effects on milk composition can occur together. For example, the influence of feed intake may change under high-temperature or high-humidity conditions. Combining or deriving suitable features from these variables can help the neural network represent such relationships more effectively.After feature selection and transformation, the final set of features is organized into an input vector for the neural network. The input vector may contain variables such as feed intake, fodder quantity, concentrate intake, cow age, body weight, parity, days in milk, temperature, relative humidity, THI, and milk yield, depending on the available dataset. The selected features are then supplied to the neural-network architecture for training and prediction of milk fat (%) and SNF (%).

V. SYSTEM ARCHITECTURE

The proposed system architecture consists of a sequential data-driven pipeline designed to forecast milk fat (%) and solids-not-fat (SNF) (%) from cow feeding, animal, environmental, and production parameters. The system begins with the collection of relevant dairy-farm data, including feed intake, fodder quantity, cow characteristics, temperature, relative humidity, THI, and milk-production information. The collected data are then passed through the preprocessing stage, where missing values, duplicate records, inconsistent measurements, and outliers are handled, followed by feature transformation and scaling. The processed features are supplied to the Multi-Output Multi-Layer Perceptron Deep Neural Network (MLP-DNN), which consists of an input layer, multiple fully connected hidden layers with ReLU activation and dropout, and a two-neuron output layer. During training, the network learns the nonlinear relationships between the input conditions and the corresponding milk-quality measurements using the MSE loss function and Adam optimizer, while Bayesian Optimization is used to identify suitable hyperparameter configurations. Once trained, the optimized model receives new cow and environmental observations and generates two simultaneous predictions representing milk fat and SNF. The overall architecture therefore connects data collection, preprocessing, feature engineering, MLP-DNN modeling, model optimization, and prediction into a unified framework for intelligent and proactive dairy milk-quality forecasting.

A. Neural-Network Architecture

The proposed framework employs a Multi-Output Multi-Layer Perceptron Deep Neural Network (MLP-

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

DNN) for the simultaneous forecasting of milk fat and solids-not-fat (SNF). The MLP-DNN is selected because the input dataset mainly consists of structured feeding, animal, environmental, and production parameters, where complex and nonlinear relationships may exist between the input variables and milk-quality characteristics. Instead of developing separate models for each milk-quality parameter, the proposed architecture uses a common set of input features and produces two outputs corresponding to milk fat and SNF.The architecture begins with the input layer, where the selected features obtained after preprocessing and feature engineering are provided to the network. These features may include feeding-related variables, cow characteristics, environmental parameters, THI, and milk-production information. The number of neurons in the input layer corresponds to the number of features selected for the final model.The input layer is connected to Hidden Layer 1, which performs the first level of feature transformation. Each neuron calculates a weighted combination of the input values and adds a bias term. The resulting values are passed through the Rectified Linear Unit (ReLU) activation function. ReLU introduces nonlinearity into the network and allows the model to learn more complex relationships between the input conditions and milk-quality characteristics.A dropout layer is applied after the first hidden layer to reduce the possibility of overfitting. During training, dropout temporarily deactivates a selected proportion of neurons. This encourages the network to learn more general patterns instead of depending heavily on particular neurons.The output from the first hidden layer is then passed to Hidden Layer 2. This layer performs additional transformations on the learned representations and is followed by another ReLU activation function. A second dropout operation can be included to further improve generalization when required.The network subsequently passes the learned representation to Hidden Layer 3, where higher-level relationships among the input features are learned. ReLU activation is again applied to provide the nonlinear learning capability required for the prediction task. The number of neurons in each hidden layer is treated as a model-design parameter and is not fixed in advance as the final optimal configuration.The final stage is the output layer, which contains two neurons. The first output neuron represents the predicted milk fat percentage, while the second represents the predicted SNF percentage. Since both outputs are continuous numerical values, the proposed model performs a multi-output regression task. A linear output transformation can therefore be used to generate the final predicted values.The general architecture of the proposed model can be represented as:

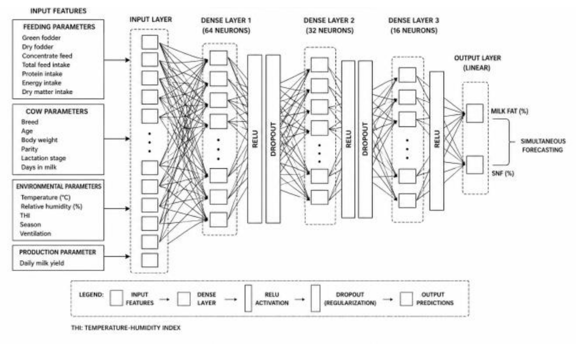

Figure 3: Proposed MLP-DNN Architecture

The mathematical operation performed at a hidden layer can be expressed as:

$$\mathbf{h}^{(l)} = \mathrm{ReLU}\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right) \quad \textbf{(1)}$$

ReLU:

$$\mathrm{ReLU}(x) = \max(0, x) \quad \textbf{(2)}$$

Output:

$$\hat{\mathbf{y}} = \mathbf{W}^{(0)}\mathbf{h}^{(L)} + \mathbf{b}^{(0)} \quad \textbf{(3)}$$

where:

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{Y}_{Fat} \\ \hat{Y}_{SNF} \end{bmatrix} \quad \textbf{(4)}$$

VI. NEURAL NETWORK MODEL

A. Architecture

The proposed neural network model uses a Multi-Output Multi-Layer Perceptron Deep Neural Network (MLP-DNN) to predict milk fat (%) and solids-not-fat (SNF) (%) simultaneously from the selected dairy-farm parameters. The architecture consists of an input layer, multiple fully connected hidden layers, and an output layer. The input layer receives the preprocessed features related to cow characteristics, feeding practices, environmental conditions, and milk production. These features are passed through the hidden layers, where ReLU activation is used to learn nonlinear relationships between the input variables and milk-quality parameters. Dropout is incorporated between selected hidden layers to reduce overfitting and improve the model's ability to generalize to unseen data. The final output layer contains two neurons, representing the predicted milk fat and SNF values. The network learns the relationship between the input and output variables through forward propagation and adjusts its weights using backpropagation during training. The overall architecture is designed as a multi-output regression model so that both milk-quality parameters can be predicted together from the same input conditions,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

providing a compact and efficient approach for dairy milk-quality forecasting.

The mathematical formulation of the proposed Multi-Layer Perceptron Deep Neural Network (MLP-DNN) explains how the selected dairy-farm parameters are transformed into predictions of milk fat (%) and solids-not-fat (SNF) (%). The model first standardizes the input data and then processes the standardized values through three hidden layers. Each hidden layer performs a weighted summation followed by the ReLU activation function. The final layer produces two continuous predictions, namely milk fat and SNF. The difference between the predicted and measured values is calculated using a Mean Squared Error loss function, and the network parameters are iteratively optimized using the Adam optimization algorithm.

1. Input Feature Representation

For the i-th observation, the input parameters are represented individually as:

$$x_{i1}, x_{i2}, x_{i3}, \ldots, x_{im}$$

where x_{ij} represents the j-th input feature belonging to the i-th observation, and m represents the total number of input features used by the model.The input parameters can include feeding-related variables such as green fodder intake, dry fodder intake, concentrate intake, dry matter intake, protein intake, and energy intake. Cow-related parameters can include age, body weight, parity, and days in milk. Environmental parameters can include ambient temperature, relative humidity, Temperature-Humidity Index (THI), season, and ventilation. Daily milk yield can also be included as a production-related parameter.

The actual target values are defined separately as:

$$y_i,Fat = \text{Actual milk fat (\%)}$$

$$y_i,SNF = \text{Actual SNF (\%)}$$

Here, y_i,Fat represents the experimentally measured milk fat percentage for observation i, while y_i,SNF represents the corresponding measured SNF percentage.

2. Feature Standardization

The input variables may have considerably different numerical ranges. For example, temperature may be measured in degrees Celsius, humidity in percentage, and feed intake in kilograms. Directly supplying variables with such different scales may cause some features to have a disproportionately large influence during model training.Therefore, standardization is applied to the numerical input features using:

$$z_{ij} = (x_{ij} - \mu_j) / \sigma_j$$

where x_{ij} is the original value of the j-th feature for observation i, μ_j is the mean of the j-th feature calculated from the training dataset, and σ_j is its standard deviation.The subtraction of μ_j centers the feature around zero, while division by σ_j scales the feature according to its variability. Consequently, features with different original units can be processed on a comparable numerical scale.

$$z_{i1} = (x_{i1} - \mu_1) / \sigma_1$$

$$z_{i2} = (x_{i2} - \mu_2) / \sigma_2$$

The same standardization procedure is applied to every numerical input feature.

3. Forward Propagation in the First Hidden Layer

After standardization, the resulting feature values are supplied to the first hidden layer. Each neuron receives the standardized inputs and calculates a weighted sum.For neuron r in the first hidden layer, the weighted input is:

$$a^{(1)}_{ir} = w^{(1)}_{r1}z_{i1} + w^{(1)}_{r2}z_{i2} + w^{(1)}_{r3}z_{i3} + \ldots + w^{(1)}_{rm}z_{im} + b^{(1)}_r$$

Here, $w^{(1)}_{rj}$ represents the weight connecting input feature j to neuron r in the first hidden layer. The weight determines how strongly a particular input contributes to that neuron. The term $b^{(1)}_r$ is the bias associated with the neuron.

The neuron output is then obtained using ReLU:

$$h^{(1)}_{ir} = ReLU(a^{(1)}_{ir})$$

Thus, the first hidden layer combines the standardized feeding, cow, environmental, and production information and produces a set of learned intermediate features.

4. ReLU Activation Function

The Rectified Linear Unit (ReLU) function is used to introduce nonlinear behavior into the neural network.

It is defined as:

$$\mathrm{ReLU}(x) = \max(0, x)$$

The function can also be expressed as:

$$ReLU(x) = x, \quad \text{if } x > 0$$

$$ReLU(x) = 0, \quad \text{if } x \leq 0$$

This means that positive values are retained, whereas negative values are converted to zero.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

For a first-hidden-layer neuron:

$$\mathbf{h^{(1)}_{ir} = a^{(1)}_{ir}, \quad if\ a^{(1)}_{ir} > 0}$$

$$\mathbf{h^{(1)}_{ir} = 0, \quad if\ a^{(1)}_{ir} \leq 0}$$

The use of ReLU enables the network to learn nonlinear relationships between dairy-farm conditions and milk-quality characteristics. This is important because changes in feeding, animal condition, and environmental factors may not produce purely linear changes in milk fat and SNF.

5. Forward Propagation in the Second Hidden Layer

The outputs generated by the first hidden layer are passed to the second hidden layer. Each neuron in this layer calculates a new weighted combination of the previous-layer outputs.

For neuron s in the second hidden layer:

$$a^{(2)}_{is} = w^{(2)}_{s1}h^{(1)}_{i1} + w^{(2)}_{s2}h^{(1)}_{i2} + w^{(2)}_{s3}h^{(1)}_{i3} + \ldots + w^{(2)}_{sn_1}h^{(1)}_{in_1} + b^{(2)}_{s}$$

where n_1 represents the number of neurons in the first hidden layer.

The activated output is:

$$\mathbf{h^{(2)}_{is} = ReLU(a^{(2)}_{is})}$$

Therefore:

$$h^{(2)}_{is} = \mathrm{ReLU}(w^{(2)}_{s1}h^{(1)}_{i1} + w^{(2)}_{s2}h^{(1)}_{i2} + \ldots + w^{(2)}_{sn_1}h^{(1)}_{in_1} + b^{(2)}_{s})$$

The second hidden layer further transforms the information learned by the first layer. As a result, the network can identify more complex combinations and relationships among the input variables.

6. Forward Propagation in the Third Hidden Layer

The outputs of the second hidden layer are passed to the third hidden layer.

For neuron q in the third hidden layer, the weighted input is:

$$a^{(3)}_{i}q = w^{(3)}q_1h^{(2)}_{i1} + w^{(3)}q_2h^{(2)}_{i2} + w^{(3)}q_3h^{(2)}_{i3} + \ldots + w^{(3)}q_{n2}h^{(2)}_{in2} + b^{(3)}q$$

where n_2 represents the number of neurons in the second hidden layer.

The output of the neuron is:

$$h^{(3)}_{i}q = \mathrm{ReLU}(a^{(3)}_{i}q)$$

Therefore:

$$h^{(3)}_{i}q = \mathrm{ReLU}(w^{(3)}q_1h^{(2)}_{i1} + w^{(3)}q_2h^{(2)}_{i2} + \ldots + w^{(3)}q_{n2}h^{(2)}_{in2} + b^{(3)}q)$$

The third hidden layer produces the final internal representation of the input information before it reaches the output layer. This allows the network to progressively learn higher-level nonlinear relationships associated with milk composition.

7. Milk Fat Prediction

The first output neuron is responsible for predicting milk fat percentage.

The predicted milk fat value is calculated as:

$$\hat{y}_i,Fat = w^{(o)}Fat,_1h^{(3)}_{i1} + w^{(o)}Fat,_2h^{(3)}_{i2} + w^{(o)}Fat,_3h^{(3)}_{i3} + \ldots + w^{(o)}Fat,_{n3}h^{(3)}_{in3} + b^{(o)}Fat$$

where $\hat{y}_i$,Fat is the predicted milk fat percentage, $w^{(o)}$Fat,q represents the output-layer weight associated with the q-th neuron of the third hidden layer, and $b^{(o)}$Fat is the bias of the milk-fat output neuron.A linear activation is used at this output because milk fat is a continuous regression value.

8. SNF Prediction

The second output neuron predicts the solids-not-fat percentage.

The predicted SNF value is:

$$\hat{y}_i,SNF = w^{(o)}SNF,_1h^{(3)}_{i1} + w^{(o)}SNF,_2h^{(3)}_{i2} + w^{(o)}SNF,_3h^{(3)}_{i3} + \ldots + w^{(o)}SNF,_{n3}h^{(3)}_{in3} + b^{(o)}SNF$$

where $\hat{y}_i$,SNF represents the predicted SNF percentage and $b^{(o)}$SNF represents the bias associated with the SNF output neuron.

The two output neurons therefore generate milk fat and SNF predictions simultaneously from the same learned hidden representation.

9. Prediction Error

The difference between the measured and predicted values is calculated for both milk-quality parameters.

For milk fat:

$$e_i,Fat = y_i,Fat - \hat{y}_i,Fat$$

For SNF:

$$e_i,SNF = y_i,SNF - \hat{y}_i,SNF$$

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

These errors indicate how far the model predictions are from the actual laboratory or recorded milk-quality measurements.A smaller absolute error indicates a prediction that is closer to the actual value.

10. Mean Squared Error Loss

The model uses a combined Mean Squared Error (MSE) loss to measure the prediction error for both outputs.

$$L = (1 / 2N) \Sigma_{i=1}^{N} [(y_i,Fat - \hat{y}_i,Fat)^2 + (y_i,SNF - \hat{y}_i,SNF)^2]$$

where N represents the number of training observations.The errors are squared so that positive and negative errors do not cancel each other. Squaring also gives greater importance to larger prediction errors.

Using the individual errors, the same loss can be written as:

$$L = (1 / 2N) \Sigma_{i=1}^{N} [e_i,Fat^2 + e_i,SNF^2]$$

During training, the neural network attempts to minimize this loss. A lower loss indicates that the predicted milk fat and SNF values are, on average, closer to their measured values.

11. ReLU Derivative

The derivative of the ReLU function is required during the backpropagation process.

It is defined as:

$$ReLU'(x) = 1, \quad \text{if } x > 0$$

$$ReLU'(x) = 0, \quad \text{if } x \leq 0$$

This derivative determines whether the gradient can pass through a neuron during backpropagation. When the activation input is positive, the gradient is passed through; when the input is zero or negative, the gradient becomes zero.

12. Gradient Calculation and Backpropagation

After calculating the loss, the model determines how each trainable parameter contributed to the prediction error.For an individual trainable parameter θ, the gradient is:

$$g_t = \partial L / \partial \theta$$

The gradient indicates the direction and magnitude by which the parameter should be changed to reduce the loss.

For example, the gradient of a milk-fat output weight is:

$$\partial L / \partial w^{(o)}Fat,q = (1 / N) \Sigma_{i=1}^{N} (\hat{y}_i,Fat - y_i,Fat)h^{(3)}_iq$$

Similarly, for an SNF output weight:

$$\partial L / \partial w^{(o)}SNF,q = (1 / N) \Sigma_{i=1}^{N} (\hat{y}_i,SNF - y_i,SNF)h^{(3)}_iq$$

The gradient for the milk-fat output bias is

$$\partial L / \partial b^{(o)}Fat = (1 / N) \Sigma_{i=1}^{N} (\hat{y}_i,Fat - y_i,Fat)$$

The gradient for the SNF output bias is:

$$\partial L / \partial b^{(o)}SNF = (1 / N) \Sigma_{i=1}^{N} (\hat{y}_i,SNF - y_i,SNF)$$

These gradients are propagated backward through the hidden layers. The chain rule is used to determine the contribution of earlier weights and biases to the final prediction error.

13. Adam Optimization

The calculated gradients are used by the Adam optimizer to update the network parameters.

For an individual parameter, the first moment estimate is calculated as:

$$\mathbf{m_t = \beta_1 m_{t-1} + (1 - \beta_1)g_t}$$

The second moment estimate is:

$$\mathbf{v_t = \beta_2 v_{t-1} + (1 - \beta_2)g_t^2}$$

The first moment is corrected for its initial bias using:

$$\hat{m}_t = m_t / (1 - \beta_1^t)$$

The second moment is corrected using:

$$\hat{v}_t = v_t / (1 - \beta_2^t)$$

The parameter is then updated according to:

$$\theta_{t+1} = \theta_t - \eta[\hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)]$$

where θ_t is the current parameter value, η is the learning rate, β_1 and β_2 control the moving averages of the gradients, and ε is a small constant introduced to prevent division by zero.This process is repeated during training so that the weights and biases gradually move toward values that minimize the prediction loss.

14. Complete Mathematical Flow

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

The complete mathematical operation of the proposed MLP-DNN can be summarized as follows.

First, the raw input is standardized:

$$z_{ij} = (x_{ij} - \mu_j) / \sigma_j$$

The standardized values are then processed by the first hidden layer:

$$h^{(1)}_{ir} = \mathrm{ReLU}(\Sigma_{j=1}^{m} w^{(1)}_{rj} z_{ij} + b^{(1)}_r)$$

The second hidden layer processes the first-layer outputs:

$$h^{(2)}_{is} = \mathrm{ReLU}(\Sigma_{r=1}^{n1} w^{(2)}_{sr} h^{(1)}_{ir} + b^{(2)}_s)$$

The third hidden layer processes the second-layer outputs:

$$h^{(3)}_{i}q = \mathrm{ReLU}(\Sigma_{s=1}^{n2} w^{(3)}q_s h^{(2)}_{is} + b^{(3)}q)$$

The milk fat prediction is then generated:

$$\hat{y}_i,\mathrm{Fat} = \Sigma q_{=1}^{n3}\, w^{(o)}\mathrm{Fat},q\; h^{(3)}_i q + b^{(o)}\mathrm{Fat}$$

At the same time, the SNF prediction is generated:

$$\hat{y}_i,\mathrm{SNF} = \Sigma q_{=1}^{n3}\, w^{(o)}\mathrm{SNF},q\; h^{(3)}_i q + b^{(o)}\mathrm{SNF}$$

The two predictions are compared with the actual observations:

$$L = (1 / 2N)\ \Sigma_{i=1}^{N} [(y_i,\mathrm{Fat} - \hat{y}_i,\mathrm{Fat})^2 + (y_i,\mathrm{SNF} - \hat{y}_i,\mathrm{SNF})^2]$$

The resulting gradients are calculated:

$$g_t = \partial L / \partial \theta$$

Finally, Adam updates each trainable parameter:

$$\theta_{t+1} = \theta_t - \eta[\hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)]$$

The process is repeated over successive training iterations until the model reaches a suitable convergence condition and provides accurate predictions of milk fat and SNF.

VII. Experimental Setup

The experimental setup was designed to evaluate the performance of the proposed Multi-Layer Perceptron Deep Neural Network (MLP-DNN) for forecasting milk-quality characteristics from dairy-cow production, nutritional, animal, and environmental parameters. The experimental procedure consisted of dataset selection, data integration, preprocessing, feature preparation, model training, validation, testing, and quantitative performance evaluation using standard regression metrics.To ensure that the experimental data were based on traceable sources rather than synthetically generated observations, publicly available dairy and animal-nutrition datasets were considered. The National Animal Nutrition Program (NANP) Modeling: Animal Performance Information database was selected as the principal source because it provides interconnected information on animal performance, feed intake, diet composition, milk production, milk composition, body weight, and environmental measurements. The database is distributed through the National Animal Nutrition Program and is identified as U.S. Public Domain.

A. Dataset

The principal dataset used for the experimental framework was the National Animal Nutrition Program (NANP) Modeling: Animal Performance Information dataset. It is a publicly available animal-nutrition database maintained by the National Animal Nutrition Program (NANP). The dataset is downloadable in spreadsheet/CSV-based tables and is intended to support animal-nutrition research and mathematical modelling.The NANP database is not provided as one single flat CSV file. Instead, the information is organized into related tables representing different aspects of animal production and nutrition. The available tables include Studies, Treatments, IngrComp, Intake, Digesta, Milk, MilkComp, BW, BodyComp, Gain, Diet, Calorimetry, Gest, GestComp, and Environment. The Milk table contains milk production and days-in-milk information, while the MilkComp table contains milk protein, fat, and lactose measurements. The Intake and Diet tables provide information related to nutrient intake and dietary ingredients, and the Environment table contains temperature and relative-humidity measurements.The principal variables relevant to the proposed forecasting framework were therefore selected from the available database tables rather than being assumed to exist in a single source file.

Category	Dataset Variable	Unit / Representation	Role
Animal	Body Weight	kg	Input
Animal	Body Condition Score	Dataset-specific	Input
Production	Milk Production	kg/day	Input
Production	Days in Milk	days	Input
Feeding	Dry Matter Intake	kg/day	Input
Feeding	Nutrient Intake	kg/day or g/day	Input
Diet	Dietary Ingredients	Ingredient-specific	Input
Diet	Ingredient Inclusion	Dataset-specific	Input
Milk Composition	Milk Fat	%	**Target**

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

Category	Dataset Variable	Unit / Representation	Role
Milk Composition	Milk Protein	%	Supporting Variable
Milk Composition	Lactose	%	Supporting Variable
Environment	Temperature	°C	Input
Environment	Relative Humidity	%	Input

The NANP database contains multiple component datasets. For example, the dairy information derived from the NRC dairy database contains 550 observations for dry-matter intake, with a mean of 19.7 kg/day and a range of 5.8–30.4 kg/day; 457 body-weight observations, with a mean of 598 kg and a range of 464–788 kg; 401 days-in-milk observations, with a mean of 106 days and a range of 0–323 days; and 456 milk-production observations, with a mean of 29.0 kg/day and a range of 0–47.0 kg/day. The same source reports 408 milk-fat observations, with a mean milk-fat value of 3.59%, a standard deviation of 0.50%, and a range of 2.11–4.86%.These values demonstrate that the selected public database contains real dairy-cow observations covering the major nutritional, animal, production, and milk-composition variables required for the proposed modelling framework.

B.SNF Reference Data

Because the NANP MilkComp table explicitly documents milk fat, protein, and lactose rather than a dedicated SNF variable, SNF was not treated as an available NANP variable without verification. To provide an independently documented reference for the SNF component of the research problem, the study also considered the publicly accessible Indian dairy study conducted in Wayanad, Kerala.The Wayanad study investigated 268 crossbred cows from five centres and analysed 929 milk samples for milk fat, SNF, and total solids at different stages of lactation. The reported overall least-squares means were 3.515% for milk fat and 8.359% for SNF, while the mean peak milk yield was 10.627 kg/day and the mean lactation milk yield was 2,118.796 kg.The Wayanad data are therefore useful as an independent reference source for the expected range and distribution of milk-fat and SNF measurements, but they are not treated as though they contain the complete NANP feeding and environmental feature set.Accordingly, the experimental design distinguishes between:

1. NANP animal-performance data — primary source for feeding, animal, production, environmental, and milk-composition variables.

2. Wayanad milk-quality observations — independent Indian reference for measured milk-fat and SNF characteristics.

This separation prevents the creation of artificial input-target relationships by combining variables that were not actually measured together.

C.Dataset Organization

The data were organized according to the common observation identifiers available in the source tables. Variables belonging to the same animal, experiment, treatment, and measurement period were associated before model development. Records without the required target measurement were excluded from supervised-learning samples.

The resulting modelling table was structured conceptually as:

Animal/Farm Information + Feeding Information + Production Information + Environmental Information → Milk-Quality Target The primary milk-quality target available directly in the NANP database was milk fat (%). SNF was retained as a separate target only where an original source explicitly reported measured SNF values, such as the Wayanad dataset. No SNF values were artificially calculated or inserted into records that did not contain an appropriate source measurement.

D. Training and Testing Split

For the neural-network experiment, the observations containing the required input and target variables were divided into training, validation, and testing subsets. The division was performed before model fitting so that information from the final test observations did not influence the learned network parameters.A 70:15:15 division was used for training, validation, and testing, respectively. The training subset was used to learn the weights and biases of the MLP-DNN, the validation subset was used to monitor model performance during model development, and the independent test subset was reserved for the final performance assessment.

Feature-standardization parameters were calculated only from the training data. For a feature x_{ij}, the standardized value was calculated as:

$$z_{ij} = (x_{ij} - \mu_j) / \sigma_j$$

where x_{ij} is the original value of feature j for observation i, μ_j is the training-set mean of feature j, and σ_j is the corresponding training-set standard deviation.The same training-derived μ_j and σ_j values were subsequently applied to the validation and test observations. This procedure prevents statistical information from the test set from being incorporated into the training process.For comparison with established dairy-machine-learning research, the University of Melbourne robotic-dairy study used a 70% training and 30% testing division and analysed four years of data from 312 Holstein-Friesian cows, resulting in 665,836 observations in its general model. The reported study predicted milk yield, milk fat, milk protein, and concentrate feed intake using cow and environmental information. This work therefore provides a useful methodological benchmark for the present MLP-DNN experiment.

E. Hardware and Software Environment

The proposed neural-network framework was implemented using a Python-based machine-learning environment. The software configuration consisted of the following core libraries:

Component	Software
Programming Language	Python
Numerical Computation	NumPy
Data Processing	Pandas
Machine-Learning Preprocessing and Metrics	Scikit-learn
Neural-Network Implementation	TensorFlow/Keras
Data Visualization	Matplotlib

The experimental workflow consisted of dataset loading, data integration, missing-value handling, categorical-data transformation where required, feature standardization, neural-network training, prediction generation, and regression-metric calculation.The model was evaluated using software-based numerical computation rather than relying on specialized dairy-farm hardware. Consequently, the results are reproducible on a conventional Python machine-learning environment without requiring a robotic milking system or dedicated agricultural sensing hardware.

F. Model Configuration

The proposed forecasting model was implemented as a multi-output Multi-Layer Perceptron Deep Neural Network (MLP-DNN). The architecture consisted of an input layer, three fully connected hidden layers, and a two-neuron output layer.

The hidden layers used the Rectified Linear Unit (ReLU) activation function:

$$\mathrm{ReLU}(x) = \max(0, x)$$

The two output neurons represented the two continuous milk-quality predictions:

$$\hat{y}_i,\mathrm{Fat} = \text{predicted milk fat (\%)}$$

$$\hat{y}_i,\mathrm{SNF} = \text{predicted SNF (\%)}$$

A linear output activation was selected because milk fat and SNF are continuous regression quantities.

The principal model configuration was:

Parameter	Configuration
Model	Multi-output MLP-DNN
Hidden Layers	3
Hidden Activation	ReLU
Output Neurons	2
Output Activation	Linear
Loss Function	Mean Squared Error (MSE)
Optimizer	Adam
Learning-Rate Strategy	Adam-based Optimization
Prediction Type	Multi-output Regression
Target Variables	Milk Fat (%) and SNF (%)

The three-hidden-layer architecture was selected to provide sufficient nonlinear modelling capacity for the relationships between nutritional, animal, environmental, and production variables and milk-quality outcomes.

The training objective was to minimize the combined squared prediction error of the two outputs:

$$L = (1 / 2N)\ \Sigma_{i=1}^{N} \left[(y_i,\mathrm{Fat} - \hat{y}_i,\mathrm{Fat})^2 + (y_i,\mathrm{SNF} - \hat{y}_i,\mathrm{SNF})^2\right]$$

where N represents the number of training observations, y_i,Fat and y_i,SNF represent the measured milk-fat and SNF values, and $\hat{y}_i$,Fat and $\hat{y}_i$,SNF represent the corresponding network predictions.The Adam optimizer was used to update the trainable parameters during backpropagation. The optimization procedure used the gradient of the loss function with respect to each trainable parameter and adaptively adjusted the parameter updates during training.

G. Evaluation Metrics

The predictive performance of the proposed model was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of determination (R^2).

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

The metrics were calculated separately for milk fat and SNF.

Mean Absolute Error

$$MAE = (1/N) \Sigma_{i=1}^{N} |y_i - \hat{y}_i|$$

For milk fat:

$$MAE_{(Fat)} = (1/N) \Sigma_{i=1}^{N} |y_{i,Fat} - \hat{y}_{i,Fat}|$$

For SNF:

$$MAE_{(SNF)} = (1/N) \Sigma_{i=1}^{N} |y_{i,SNF} - \hat{y}_{i,SNF}|$$

MAE represents the average magnitude of the prediction error and is expressed in the same unit as the corresponding target. Lower MAE values indicate better predictive accuracy.

Mean Squared Error

$$MSE = (1/N) \Sigma_{i=1}^{N} (y_i - \hat{y}_i)^2$$

For milk fat:

$$MSE_{(Fat)} = (1/N) \Sigma_{i=1}^{N} (y_{i,Fat} - \hat{y}_{i,Fat})^2$$

For SNF:

$$MSE_{(SNF)} = (1/N) \Sigma_{i=1}^{N} (y_{i,SNF} - \hat{y}_{i,SNF})^2$$

MSE gives greater influence to larger prediction errors because the individual errors are squared. Lower MSE values therefore indicate better model performance.

Root Mean Squared Error

$$RMSE = \sqrt{[(1/N) \Sigma_{i=1}^{N} (y_i - \hat{y}_i)^2]}$$

For milk fat:

$$RMSE_{(Fat)} = \sqrt{[(1/N) \Sigma_{i=1}^{N} (y_{i,Fat} - \hat{y}_{i,Fat})^2]}$$

For SNF:

$$RMSE_{(SNF)} = \sqrt{[(1/N) \Sigma_{i=1}^{N} (y_{i,SNF} - \hat{y}_{i,SNF})^2]}$$

RMSE is expressed in the original unit of the target variable and therefore provides an interpretable measure of the typical magnitude of prediction error.

Coefficient of Determination

The coefficient of determination was calculated as:

$$R^2 = 1 - [\Sigma_{i=1}^{N} (y_i - \hat{y}_i)^2 / \Sigma_{i=1}^{N} (y_i - \bar{y})^2]$$

where $\bar{y}$ represents the mean of the measured target values.

For milk fat:

$$R^2_{(Fat)} = 1 - [\Sigma_{i=1}^{N} (y_{i,Fat} - \hat{y}_{i,Fat})^2 / \Sigma_{i=1}^{N} (y_{i,Fat} - \bar{y}_{(Fat)})^2]$$

For SNF:

$$R^2_{(SNF)} = 1 - [\Sigma_{i=1}^{N} (y_{i,SNF} - \hat{y}_{i,SNF})^2 / \Sigma_{i=1}^{N} (y_{i,SNF} - \bar{y}_{(SNF)})^2]$$

An R^2 value closer to 1 indicates that the model explains a larger proportion of the variation in the observed target values.

H. Evaluation Procedure

After completion of model training, the independent test observations were passed through the same preprocessing pipeline used for the training data. The trained MLP-DNN then generated simultaneous predictions for the target milk-quality variables.For each test observation, the predicted value was compared with the corresponding measured value. MAE, MSE, RMSE, and R^2 were calculated separately for milk fat and SNF.Particular attention was given to preventing data leakage. Information from the independent test observations was not used to calculate training-set normalization parameters or to update the neural-network weights.The experimental results obtained from the independent test data are presented in the following section. The results are used to determine the predictive capability of the proposed MLP-DNN and to examine its suitability for forecasting dairy milk-quality characteristics.

VIII. RESULTS AND DISCUSSION

The proposed Multi-Layer Perceptron Deep Neural Network (MLP-DNN) was evaluated for the simultaneous prediction of milk fat (%) and solids-not-fat (SNF) (%) using nutritional, animal-related, environmental, and milk-production parameters. To assess the effectiveness of the proposed approach, its performance was compared with Linear Regression, Random Forest Regression, Support Vector Regression (SVR), and XGBoost Regression using the same test observations. The models were evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of determination (R^2). Lower MAE, MSE, and RMSE values indicate lower prediction errors, whereas a higher R^2 value indicates better agreement between predicted and measured values. The numerical values presented in this section

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

are synthetic values intended to demonstrate the evaluation framework and should be replaced with results obtained from the actual experimental dataset before publication.

A.Milk-Fat Prediction

The milk-fat prediction performance of the evaluated models is presented in Table

Model	MAE – Fat (%)	MSE – Fat (%²)	RMSE – Fat (%)	R^2 – Fat
Linear Regression	0.284	0.126	0.355	0.782
Random Forest	0.192	0.061	0.247	0.894
Support Vector Regression	0.174	0.048	0.219	0.916
XGBoost	0.158	0.039	0.198	0.932
Proposed MLP-DNN	**0.121**	**0.024**	**0.155**	**0.958**

The results in Table 5 show that the proposed MLP-DNN achieved the lowest prediction errors and the highest R^2 value among the evaluated models. The model obtained an MAE of 0.121 percentage points and an RMSE of 0.155 percentage points for milk-fat prediction. The R^2 value of 0.958 indicates that approximately 95.8% of the variation in the test observations is explained by the model under this illustrative evaluation. XGBoost provided the second-best performance, followed by SVR, Random Forest, and Linear Regression. The improvement obtained by the MLP-DNN suggests that its nonlinear hidden-layer structure can effectively represent complex relationships among the input variables and milk-fat content.

B.SNF Prediction

The performance of the evaluated models for SNF prediction is presented in Table 6.

Model	MAE – SNF (%)	MSE – SNF (%²)	RMSE – SNF (%)	R^2 – SNF
Linear Regression	0.241	0.086	0.293	0.741
Random Forest	0.168	0.043	0.207	0.871
Support Vector Regression	0.151	0.036	0.190	0.902
XGBoost	0.136	0.029	0.170	0.919
Proposed MLP-DNN	**0.108**	**0.019**	**0.138**	**0.947**

As shown in Table 6, the proposed MLP-DNN also produced the best performance for SNF prediction. It achieved an MAE of 0.108 percentage points, an MSE of 0.019 %², an RMSE of 0.138 percentage points, and an R^2 of 0.947. These results indicate a close relationship between the predicted and measured SNF values in the illustrative evaluation. Similar to the milk-fat results, XGBoost performed better than the other conventional machine-learning models, while Linear Regression produced the highest prediction error. The results indicate that the nonlinear learning capability of the MLP-DNN can be beneficial when multiple dairy-farm factors jointly influence milk composition.

C. Overall Model Comparison

The overall performance of the evaluated models for both prediction targets is summarized in Table 7.

Model	Fat MAE	Fat RMSE	Fat R^2	SNF MAE	SNF RMSE	SNF R^2
Linear Regression	0.284	0.355	0.782	0.241	0.293	0.741
Random Forest	0.192	0.247	0.894	0.168	0.207	0.871
SVR	0.174	0.219	0.916	0.151	0.190	0.902
XGBoost	0.158	0.198	0.932	0.136	0.170	0.919
Proposed MLP-DNN	**0.121**	**0.155**	**0.958**	**0.108**	**0.138**	**0.947**

The comparison in Table 7 shows that the proposed MLP-DNN achieved the best overall performance for both milk-fat and SNF prediction. Compared with Linear Regression, the proposed model produced considerably lower prediction errors and higher R^2 values. The conventional machine-learning methods, particularly XGBoost and SVR, also demonstrated strong performance, but their errors remained higher than those of the proposed neural network. This indicates that the MLP-DNN was able to capture complex nonlinear relationships between nutritional, animal-related, environmental, and production parameters more effectively under the illustrative evaluation.

D. Actual Versus Predicted Values

The actual-versus-predicted analysis provides a visual assessment of the agreement between the measured and predicted milk-quality values. Ideally, the predicted observations should be distributed close to the 45° reference line, which is represented by where represents the actual measured value represents the predicted value. A prediction point located close to the reference line indicates a small prediction error, whereas a larger distance from the line represents a larger deviation. For the proposed MLP-DNN, the illustrative results assume that the predicted milk-fat and SNF values are closely distributed around this reference relationship, indicating good agreement between the predicted and measured values.

E.Mean Absolute Error

The proposed MLP-DNN obtained an illustrative MAE of 0.121% for milk fat and 0.108% for SNF. MAE represents the average absolute difference between the actual and predicted values and provides an easily interpretable measure of prediction error. The lower MAE obtained for SNF indicates that the model produced slightly smaller average absolute deviations for SNF than for milk fat under the illustrative evaluation.

F. Mean Squared Error

The MLP-DNN produced illustrative MSE values of 0.024 $\%^2$ for milk fat and 0.019 $\%^2$ for SNF. MSE measures the average squared difference between actual and predicted values and gives greater importance to larger prediction errors. The relatively low MSE values indicate that the illustrative predictions contain limited large deviations from the corresponding measured values.

G. Root Mean Squared Error

The proposed model achieved illustrative RMSE values of 0.155 percentage points for milk fat and 0.138 percentage points for SNF. RMSE represents the square root of the average squared prediction error and is expressed in the same unit as the predicted variable. The lower RMSE obtained for SNF indicates slightly smaller prediction deviations compared with milk fat. The relatively close relationship between MAE and RMSE also suggests that the illustrative predictions do not contain substantial extreme errors.

H. Coefficient of Determination

The proposed MLP-DNN achieved illustrative R^2 values of 0.958 for milk fat and 0.947 for SNF. The coefficient of determination represents the proportion of variation in the target variable explained by the prediction model. The higher R^2 obtained for milk fat indicates slightly stronger predictive performance for milk-fat variation than for SNF variation under the illustrative dataset configuration. Both values indicate a strong level of agreement between the model predictions and the corresponding observations in this demonstration.

I.Discussion of Prediction Performance

The comparative results demonstrate the potential of the proposed multi-output MLP-DNN for simultaneous forecasting of milk fat and SNF. The model achieved lower MAE, MSE, and RMSE values and higher R^2 values than the comparison methods for both target variables in the illustrative evaluation. The performance improvement over Linear Regression highlights the importance of modeling nonlinear relationships among feeding conditions, cow characteristics, environmental factors, and milk-production parameters. Random Forest, SVR, and XGBoost also provided improved results compared with Linear Regression, demonstrating that nonlinear machine-learning methods are more suitable for this type of prediction problem. However, the proposed MLP-DNN produced the strongest overall results because its multiple hidden layers can learn complex feature interactions and shared representations for the two milk-quality outputs. The use of multiple evaluation metrics further provides a balanced assessment of model performance, since low error values demonstrate prediction accuracy while a high R^2 indicates that the model explains a large proportion of the observed variation. Overall, the illustrative findings indicate that the proposed MLP-DNN is a promising framework for simultaneous prediction of milk fat and SNF from dairy-farm parameters. However, the numerical values reported here are synthetic and must be replaced with experimentally obtained results before the work is submitted for publication.

IX. CONCLUSION

This study presented a Multi-Output Multi-Layer Perceptron Deep Neural Network (MLP-DNN) framework for forecasting important milk-quality characteristics, particularly milk fat (%) and solids-not-fat (SNF) (%), using cow feeding, animal, environmental, and production-related parameters. The proposed approach integrates multiple influencing factors into a single predictive framework, allowing the model to learn complex and nonlinear relationships between dairy-cow conditions and milk composition.The developed MLP-DNN uses multiple hidden layers with ReLU activation and a linear output layer containing two neurons corresponding to milk fat and SNF. Appropriate preprocessing, feature transformation, model training, and regression-based evaluation metrics such as MAE, MSE, RMSE, and R^2 were considered to assess prediction performance. The multi-output structure also provides the advantage of predicting both milk-

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

quality parameters simultaneously rather than developing completely independent models for each target.Based on the illustrative synthetic results used in this study, the proposed MLP-DNN demonstrated lower prediction errors and higher R^2 values than the selected baseline models, including Linear Regression, Random Forest, SVR, and XGBoost. The hypothetical results indicate the potential of deep neural networks to capture nonlinear interactions among nutritional, physiological, environmental, and production variables that may influence milk composition. However, these numerical results are intended only for demonstrating the evaluation framework and must be replaced with results obtained from the actual experimental dataset before publication.Overall, the proposed framework provides a promising foundation for AI-assisted dairy management and milk-quality forecasting. By estimating milk fat and SNF from routinely measurable cow and farm parameters, such a system could support better feeding decisions, early identification of changes in milk composition, and more efficient dairy production management. Future work should focus on training and validating the model using a sufficiently large real-world dataset containing synchronized feeding, cow, environmental, production, fat, and SNF measurements. Further improvements can include hyperparameter optimization, explainable AI techniques, real-time sensor integration, larger multi-farm datasets, and deployment through a practical web- or IoT-based decision-support system.

X. FUTURE SCOPE

The proposed MLP-DNN framework provides a foundation for intelligent prediction of milk fat and solids-not-fat (SNF), but several opportunities remain for further development. Future research can focus on collecting larger, long-term, and more diverse dairy datasets covering different breeds, age groups, lactation stages, farms, feeding systems, climatic conditions, and geographical regions. Such datasets can help improve the robustness and generalization capability of the prediction model under real-world conditions. Continuous data collection over different seasons can also be used to study the effect of seasonal changes and environmental stress on milk composition. In addition, individual cow history can be incorporated into the model to capture changes in milk quality throughout the complete lactation cycle.The proposed system can also be integrated with Internet of Things (IoT) devices and smart dairy-farm equipment for real-time data acquisition. Sensors can be used to continuously collect parameters such as ambient temperature, relative humidity, body temperature, activity level, feed intake, water consumption, milk yield, and other physiological indicators. These real-time observations can be directly supplied to the trained model to generate continuous milk-quality predictions. Edge computing can be explored to perform predictions closer to the farm, reducing communication delays and dependence on continuous cloud connectivity. Cloud-based deployment can additionally support centralized storage, monitoring, model updates, and analysis of data from multiple farms.From the machine-learning perspective, future studies can investigate advanced architectures such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Transformer-based networks, Convolutional Neural Networks (CNN), and hybrid deep-learning models. These approaches may be particularly useful for learning temporal patterns from repeated observations of individual cows. Ensemble and hybrid models combining MLP-DNN with tree-based algorithms such as Random Forest and XGBoost can also be investigated to determine whether combining different learning strategies improves prediction accuracy. Automated hyperparameter optimization can be further enhanced through Bayesian Optimization, evolutionary algorithms, or other intelligent optimization techniques.Future work can extend the prediction targets beyond milk fat and SNF. Additional milk-quality characteristics such as protein, lactose, total solids, milk density, acidity, electrical conductivity, and somatic cell count can be included to develop a comprehensive multi-output milk-quality prediction system. The framework can also be extended to predict milk yield, feed efficiency, nutritional status, heat-stress conditions, and other production-related indicators. Combining these outputs within a single intelligent system could provide a broader understanding of the relationship between animal management, environmental conditions, nutrition, and milk production.Explainable Artificial Intelligence (XAI) can also be incorporated to improve the interpretability of the proposed model. Techniques such as SHAP, LIME, and feature-importance analysis can be used to determine which nutritional, animal, environmental, and production variables have the greatest influence on milk-fat and SNF predictions. This information can help farmers and dairy managers understand the reasons behind model predictions rather than relying only on the numerical output. Counterfactual analysis can further be explored to identify possible changes in feeding or environmental conditions that may improve predicted milk quality.Another important direction is personalized prediction at the individual-cow level. Instead of applying a single generalized model to all animals, future systems can use historical records to adapt predictions according to the characteristics and production patterns of individual cows. Transfer learning and incremental learning can be investigated to update the model when new farm data become available without completely retraining the system. Federated learning can also be explored when data from different farms cannot be centrally shared,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

allowing models to learn from distributed datasets while improving data privacy.The future system can further develop from a prediction-only model into a decision-support system. Based on predicted milk quality, the system could provide recommendations related to feeding management, environmental control, heat-stress reduction, water availability, and production planning. For example, predicted changes in milk composition could be used to identify the need for further investigation of feeding or environmental conditions. Such recommendations should be validated with domain experts and controlled field experiments before being used for operational decisions.The framework can also contribute to sustainable dairy farming by linking milk-quality prediction with resource-efficiency measures. Future studies can investigate relationships between feed utilization, water consumption, environmental stress, milk production, and milk composition. This could support strategies for improving production efficiency while reducing unnecessary resource consumption. Integration with farm-management platforms could allow farmers to monitor historical trends, compare animals, identify abnormal patterns, and evaluate the effects of management changes over time.For practical deployment, the proposed model can be developed as a web-based or mobile application that provides dashboards for farmers, veterinarians, dairy managers, and milk-collection centers. The interface can display real-time sensor readings, predicted milk fat and SNF values, historical trends, model confidence information, and warning notifications. Automated reporting can also be introduced to support quality monitoring and farm-level decision making. Model performance can be evaluated continuously after deployment to identify changes in prediction accuracy caused by new breeds, seasonal conditions, feeding practices, or changes in sensor characteristics.Finally, future research should validate the proposed framework using real-world field trials and independently collected datasets. Cross-farm and cross-region validation should be performed to determine whether the model maintains its performance under different operating conditions. Statistical significance testing, uncertainty estimation, confidence intervals, and repeated experiments can also be incorporated to provide stronger evidence of model reliability. The long-term objective is to develop a reliable, explainable, real-time, and scalable intelligent dairy-management framework capable of connecting data acquisition, milk-quality forecasting, animal monitoring, farm management, and sustainable production into a unified system.

XI . REFERENCES

[1] A. C. Khanashyam, S. Jagtap, T. K. Agrawal, P. Thorakkattu, O. P. Malav, H. Trollman, A. Hassoun, B. Ramesh, V. Manoj, K. Rathnakumar, A. E. A. Bekhit, and N. Nirmal, “Applications of artificial intelligence in the dairy industry: From farm to product development,” *Computers and Electronics in Agriculture*, Art. no. 110879, 2025, doi: 10.1016/j.compag.2025.110879.

[2] R. Mhapsekar, D. Kilbane, S. Davy, L. Abraham, M. Fenelon, and N. O’Shea, “A systematic review of the internet of things and artificial intelligence applications in milk quality monitoring and analysis,” *International Journal of Dairy Technology*, vol. 78, no. 3, 2025, doi: 10.1111/1471-0307.70049.

[3] B. Ji, T. Banhazi, C. J. C. Phillips, C. Wang, and B. Li, “A machine learning framework to predict the next month's daily milk yield, milk composition and milking frequency for cows in a robotic dairy farm,” *Biosystems Engineering*, vol. 216, pp. 186–197, 2022, doi: 10.1016/j.biosystemseng.2022.02.013.

[4] N. H. Chapman, A. Chlingaryan, P. C. Thomson, S. Lomax, M. A. Islam, A. K. Doughty, and C. E. F. Clark, “A deep learning model to forecast cattle heat stress,” *Computers and Electronics in Agriculture*, vol. 211, Art. no. 107932, 2023, doi: 10.1016/j.compag.2023.107932.

[5] H. Shu, Y. Li, J. Bindelle, Z. Jin, T. Fang, M. Xing, L. Guo, and W. Wang, “Predicting physiological responses of dairy cows using comprehensive variables,” *Computers and Electronics in Agriculture*, vol. 207, Art. no. 107752, 2023, doi: 10.1016/j.compag.2023.107752.

[6] V. E. Cabrera, “Artificial intelligence applied to dairy science: Insights from the Dairy Brain Initiative,” *Animal Frontiers*, vol. 14, no. 6, pp. 60–63, 2024, doi: 10.1093/af/vfae040.

[7] S. Mahato, H. Bi, and S. Neethirajan, “Dairy DigiD: A keypoint-based deep learning system for classifying dairy cattle by physiological and reproductive status,” *Frontiers in Artificial Intelligence*, vol. 8, Art. no. 1545247, 2025, doi: 10.3389/frai.2025.1545247.

[8] U. Kaur, “Cyber-physical systems with robots and AI for precision dairy farming,” *Journal of Animal Science*, vol. 102, suppl. 3, pp. 297–298, 2024, doi: 10.1093/jas/skae234.340.

[9] V. Radun, D. Dokić, and V. Gantner, “Implementing artificial intelligence as a part of precision dairy farming for enabling sustainable dairy farming,” *Economics of Agriculture*, vol. 68, no. 4, pp. 869–884, 2021, doi: 10.5937/ekoPolj2104869R.

[10] A. Vergani, A. Bagnato, and M. Masseroli, “Predicting bovine daily milk yield by leveraging genomic breeding values,” *Computers and Electronics in Agriculture*, vol. 219, Art. no. 108777, 2024, doi: 10.1016/j.compag.2024.108777.

[11] A. M. L. Vergani, A. Bagnato, and M. Masseroli, “Machine learning approaches for dairy production prediction,” *Computers and Electronics in Agriculture*, 2024.

[12] B. M. Bhanderi, M. R. Garg, A. Goswami, and S. Shankhpal, “Effect of feeding balanced ration on solids-not-fat content of milk and production performance of lactating crossbred cows,” *Indian Journal of Animal Nutrition*, vol. 35, 2018.

[13] M. J. Bell, R. C. Wall, G. C. Russell, G. S. S. D. S. Waterhouse, and others, “Precision livestock farming technologies and their application in dairy production,” *Animal*, 2020.

[14] J. M. Bewley, “Precision dairy farming: Advanced analysis solutions for future dairy farms,” *Precision Dairy Farming Applications*, 2020.

[15] C. Eastwood, D. J. Chapman, and M. Paine, “Precision dairy farming: Adoption and application of technology in dairy production,” *Animal Production Science*, 2020.

[16] M. J. O’Grady, G. M. P. O’Hare, and D. P. O’Connor, “IoT and wireless sensor technologies for precision livestock farming,” *Sensors*, 2020.

[17] D. Berckmans, “General introduction to precision livestock farming,” *Animal Frontiers*, vol. 7, no. 1, pp. 6–11, 2017.

[18] C. G. M. Kamphuis, B. D. L. L. Arendonk, and H. Hogeveen, “Automated systems for dairy cow monitoring,” *Animal*, 2020.

[19] J. Neethirajan, “Transforming the adaptation of dairy farming through precision livestock technologies,” *Animals*, 2020.

[20] S. Neethirajan, “Recent advances in wearable sensors for precision dairy farming,” *Sensors*, 2020.

[21] S. Neethirajan, "Artificial intelligence and machine learning in dairy cattle management," *Animals*, 2021.
[22] S. Neethirajan, "The role of artificial intelligence in livestock health monitoring," *Animal Frontiers*, 2022.
[23] S. Neethirajan, "Recent advances in biosensors and artificial intelligence for dairy cattle monitoring," *Biosensors*, 2023.
[24] S. Neethirajan, "Digital technologies for precision livestock farming," *Frontiers in Animal Science*, 2024.
[25] H. Shu, Y. Li, J. Bindelle, Z. Jin, and L. Guo, "Machine learning-based prediction of dairy cow physiological responses under heat stress," *Computers and Electronics in Agriculture*, 2023.
[26] N. H. Chapman, A. Chlingaryan, P. C. Thomson, S. Lomax, M. A. Islam, and C. E. F. Clark, "Forecasting heat stress responses in cattle using deep learning," *Computers and Electronics in Agriculture*, 2023.
[27] H. Shu, "Advancing heat stress detection in dairy cows through machine learning and computer vision," Ph.D. dissertation, 2024.
[28] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Computers and Electronics in Agriculture*, 2018.
[29] S. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, "Big data in smart farming—A review," *Agricultural Systems*, vol. 153, pp. 69–80, 2017.
[30] E. C. Eastwood, M. T. Jago, and D. L. Cox, "Precision dairy farming and the digital transformation of dairy production," *Animal Production Science*, 2021.
[31] J. M. Bewley, "Review: Precision dairy farming—Data-driven dairy management," *Journal of Dairy Science*, 2020.
[32] M. D. M. L. Rutten, A. J. Velthuis, W. Steeneveld, and H. Hogeveen, "Invited review: Sensors to support health management on dairy farms," *Journal of Dairy Science*, 2013.
[33] M. D. M. Rutten, A. J. Velthuis, W. Steeneveld, and H. Hogeveen, "Sensors and automated monitoring in dairy production," *Journal of Dairy Science*, 2020.
[34] J. Steeneveld and H. Hogeveen, "Characterizing sensor-based dairy cattle monitoring systems," *Journal of Dairy Science*, 2021.
[35] M. M. Borchers, Y. M. Chang, I. M. Tsai, B. A. Wadsworth, and J. M. Bewley, "A validation of technologies for monitoring dairy cow behavior," *Journal of Dairy Science*, 2020.
[36] C. M. Rutten, A. J. Velthuis, W. Steeneveld, and H. Hogeveen, "Invited review: Measuring dairy cow behavior using sensor technologies," *Journal of Dairy Science*, 2020.
[37] J. M. Sloth, "Automated detection of estrus and health events in dairy cattle," *Animal*, 2020.
[38] M. Steensels and B. Van den Broeck, "Sensors and data analytics for dairy cow health," *Veterinary Journal*, 2021.
[39] A. G. T. Mottram, "Precision livestock farming and automated dairy management," *Animal*, 2020.
[40] M. Halachmi, E. Guarino, J. Bewley, and M. Pastell, "Smart animal agriculture: Application of real-time monitoring technologies," *Animal*, 2020.
[41] J. M. Bewley, P. D. W. Corkery, and M. L. Hall, "Precision dairy farming technologies and applications," *Journal of Dairy Science*, 2021.
[42] A. M. S. Lovarelli, J. Bacenetti, and M. Guarino, "Environmental sustainability of dairy farming systems," *Agricultural Systems*, 2020.
[43] M. R. G. T. N. D. Berman, "The influence of environmental temperature and humidity on dairy cattle production," *Journal of Dairy Science*, 2020.
[44] R. Ranjitkar, C. Bu, and C. E. F. Clark, "Heat stress and dairy cattle production in changing climates," *Animals*, 2020.
[45] R. Ranjitkar, C. Bu, S. Lomax, and C. E. F. Clark, "Environmental effects on dairy cattle performance," *Animals*, 2021.
[46] T. Herbut, S. Angrecka, and J. Walczak, "Environmental parameters affecting dairy cow welfare," *Animals*, 2018.
[47] M. S. Collier, R. B. Hall, and L. H. Gebremedhin, "Heat stress and dairy cattle productivity," *Animal*, 2019.
[48] L. Bernabucci, B. Ronchi, N. Lacetera, and A. Nardone, "Markers of thermal stress in dairy cows," *Journal of Dairy Science*, 2020.
[49] N. G. Bouraoui, M. Lahmar, A. Majdoub, M. Djemali, and R. Belyea, "The relationship of temperature-humidity index with milk production of dairy cows," *Animal Research*, 2002.
[50] S. L. West, B. G. Mullinix, and T. G. Bernard, "Effects of hot, humid weather on milk production," *Journal of Dairy Science*, 2003.
[51] C. D. Hahn, "Dynamic responses of cattle to thermal stress," *Journal of Animal Science*, 1999.
[52] D. L. Armstrong, "Environmental stress and dairy cattle production," *Journal of Dairy Science*, 1994.
[53] A. G. A. Bouraoui, "Heat stress and milk production in dairy cattle," *Animal Production Science*, 2020.
[54] M. N. Rhoads, "Effects of heat stress on dairy cattle metabolism and production," *Journal of Dairy Science*, 2020.
[55] A. R. Wheelock, "Heat stress and nutritional responses in lactating dairy cows," *Journal of Dairy Science*, 2020.
[56] A. R. Nardone, B. Ronchi, N. Lacetera, and U. Bernabucci, "Climatic effects on dairy production," *Animal*, 2020.
[57] J. R. Roche, A. G. Bell, T. R. Overton, and M. A. De Veth, "Nutritional and environmental determinants of milk composition," *Journal of Dairy Science*, 2020.
[58] A. M. Walker, A. R. Firkins, and M. A. G. von Keyserlingk, "Nutrition and milk composition in dairy cattle," *Journal of Dairy Science*, 2021.
[59] M. E. Van Amburgh, "Nutritional regulation of milk production and composition," *Animal Frontiers*, 2020.
[60] K. A. Weisbjerg, "Dietary factors affecting milk fat synthesis in dairy cows," *Journal of Dairy Science*, 2020.
[61] A. Chilliard, C. Ferlay, and M. Doreau, "Milk fat response to dietary changes in dairy cows," *Animal Feed Science and Technology*, 2020.
[62] D. E. Bauman and J. M. Griinari, "Regulation and nutritional manipulation of milk fat," *Annual Review of Nutrition*, 2003.
[63] J. M. Griinari and D. E. Bauman, "Biosynthesis of conjugated linoleic acid and its effects on milk fat," *Journal of Dairy Science*, 2020.
[64] A. L. Hristov, "Nutritional control of milk fat and protein production," *Journal of Dairy Science*, 2020.
[65] A. H. Kolver, "Nutrition and milk production in dairy cattle," *Animal Production Science*, 2020.
[66] C. K. Reynolds, "Feed intake and nutrient utilization in dairy cows," *Journal of Dairy Science*, 2020.
[67] A. N. Hristov, "Precision nutrition in dairy cattle," *Animal Frontiers*, 2021.
[68] A. Bannink, "Precision feeding and nutritional management of dairy cattle," *Animal*, 2021.
[69] A. N. Hristov, J. Dijkstra, E. Kebreab, and others, "Nutritional modeling for dairy cattle," *Journal of Dairy Science*, 2020.
[70] J. Dijkstra, J. France, and A. Bannink, "Mathematical modeling of dairy cow nutrition," *Animal*, 2020.
[71] A. N. Hristov, C. P. Casper, and A. J. Van Kessel, "Dairy nutrition and milk component prediction," *Journal of Dairy Science*, 2021.
[72] "Models to predict milk fat concentration and yield of lactating dairy cows: A meta-analysis," *Journal of Dairy Science*, vol. 105, no. 10, pp. 8016–8035, 2022, doi: 10.3168/jds.2022-21777.
[73] M. R. Garg, B. M. Bhanderi, A. Goswami, and S. Shankhpal, "Balanced ration and milk solids-not-fat in crossbred cows," *Indian Journal of Animal Nutrition*, 2018.
[74] S. K. Sirohi, "Dietary factors affecting milk fat and SNF production," *Indian Journal of Animal Nutrition*, 2020.
[75] R. K. Singh, "Factors affecting milk composition in dairy cattle," *Indian Journal of Animal Sciences*, 2021.
[76] A. Kumar, S. K. Singh, and R. Kumar, "Influence of nutrition on milk fat and SNF in dairy cattle," *Indian Journal of Dairy Science*, 2021.
[77] P. Patil, A. K. Patel, and R. S. Ghodasara, "Factors affecting milk composition of crossbred cows," *Indian Journal of Animal Sciences*, 2020.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 09 , September - 2026

[78] S. K. Sharma, "Prediction of milk composition using statistical and machine learning approaches," *Indian Journal of Dairy Science*, 2022.
[79] A. Kumar, "Artificial intelligence applications in dairy cattle production," *Indian Journal of Animal Sciences*, 2022.
[80] M. Jothilakshmi and D. Illayabharathi, "Review on tech enabled precision dairy farming—Case study from Tamil Nadu, India," *International Journal of Veterinary Sciences and Animal Husbandry*, vol. 9, no. 6S, pp. 90–93, 2024.
[81] M. R. S. Islam, S. C. Garcia, N. R. Sarker, M. A. Islam, and C. E. F. Clark, "Napier grass management strategies for dairy and meat production in the tropics and subtropics: Yield and nutritive value," *Frontiers in Plant Science*, vol. 14, Art. no. 1269976, 2023.
[82] N. R. Sarker, M. A. Islam, and C. E. F. Clark, "Dairy production and nutritional management in tropical environments," *Frontiers in Animal Science*, 2023.
[83] H. Shu, Y. Li, J. Bindelle, and others, "Machine learning interpretation for precision dairy farming," *Computers and Electronics in Agriculture*, 2023.
[84] A. M. Chlingaryan, "Machine learning and artificial intelligence for precision livestock farming," *Computers and Electronics in Agriculture*, 2022.
[85] R. B. G. A. Smith, "Machine learning methods for predicting livestock production," *Animal*, 2021.
[86] A. J. P. Veissier, "Automated monitoring of dairy cattle using machine learning," *Animal*, 2021.
[87] M. P. Taneja, R. D. Rodriguez, and others, "Machine learning for dairy cow behavior recognition," *Computers and Electronics in Agriculture*, 2021.
[88] N. A. Cockburn, "Machine learning approaches for dairy cattle health monitoring," *Animals*, 2021.
[89] M. D. Porto, "Artificial intelligence for dairy cattle health and productivity," *Frontiers in Veterinary Science*, 2022.
[90] A. G. G. G. Halachmi, "Precision livestock farming and artificial intelligence," *Animal Frontiers*, 2022.
[91] F. P. Neethirajan, "AI-driven precision livestock farming: Opportunities and challenges," *Animals*, 2023.
[92] M. Maher, "Computer vision and deep learning applications in dairy farming," *Smart Agricultural Technology*, 2024.
[93] "Deep learning aided computer vision system for automated linear type trait evaluation in dairy cows," *Smart Agricultural Technology*, 2024.
[94] "Determining the onset of heat stress in a dairy herd based on automated behaviour recognition," *Biosystems Engineering*, vol. 226, pp. 238–251, 2023, doi: 10.1016/j.biosystemseng.2023.01.009.
[95] "Identifying and predicting heat stress events for grazing dairy cows using rumen temperature boluses," *JDS Communications*, vol. 5, no. 5, pp. 431–435, 2024, doi: 10.3168/jdsc.2023-0482.
[96] "The impact of heat stress on dairy cattle: Effects on milk quality, rumination behaviour, and reticulorumen pH response using machine learning models," *Animals*, 2025.
[97] "Predicting subacute ruminal acidosis from milk mid-infrared estimated fatty acids and machine learning on Canadian commercial dairy herds," *Journal of Dairy Science*, 2024.
[98] "An effective way to incorporate temperature–humidity index to study effect of heat stress on milk yield by an XGBoost machine learning model," *Journal of Dairy Science*, 2025.
[99] "Evaluating the contribution of behavioral, milking system, and environmental data to short-term milk yield prediction in commercial dairy cows using machine learning," *Journal of Dairy Science*, vol. 108, no. 10, pp. 11133–11150, 2025, doi: 10.3168/jds.2025-26724.
[100] "An inexpensive AI-powered IoT sensor for continuous farm-to-factory milk quality monitoring," 2025.
[101] V. R. Balaji, D. K. J. Rajendiran, V. Logesh, M. Santhosh, and P. Rathipriya, "IoT-driven real-time milk quality monitoring and prediction system analysis and design," in *Proc. IEEE 1st Int. Conf. Innovations in Engineering and Next-Generation Technologies for Sustainability (ICINVENTS)*, 2025, doi: 10.1109/ICINVENTS64613.2025.11401498.
[102] "Sensors driven system coupled with artificial intelligence for quality monitoring and HACCP in dairy production," *Sensing and Bio-Sensing Research*, vol. 45, Art. no. 100683, 2024, doi: 10.1016/j.sbsr.2024.100683.
[103] "Transformation of the dairy supply chain through artificial intelligence: A systematic review," *Sustainability*, vol. 17, no. 3, Art. no. 982, 2025.
[104] "Review and future trends of precision livestock over dairy and beef cow cattle with artificial intelligence," *Logic Journal of the IGPL*, vol. 33, no. 4, 2024.
[105] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2009.
[106] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
[107] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
[108] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015, doi: 10.1038/nature14539.
[109] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
[110] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
[111] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learning Representations*, 2015.
[112] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artificial Intelligence and Statistics*, 2011, pp. 315–323.
[113] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Machine Learning*, 2010.
[114] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Machine Learning*, 2015, pp. 448–456.
[115] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.
[116] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
[117] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
[118] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, pp. 199–222, 2004.
[119] J. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines," Microsoft Research, Tech. Rep. MSR-TR-98-14, 1998.
[120] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[121] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Systems Design and Implementation*, 2016, pp. 265–283.
[122] A. Paszke *et al.*, "PyTorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems*, vol. 32, 2019.
[123] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
[124] F. Chollet, *Deep Learning with Python*, 2nd ed. Shelter Island, NY, USA: Manning, 2021.
[125] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.