

# Smart GIR Cow's Disease Prediction & Supporting System Using Artificial Intelligence <sup>†</sup>

A. Vijayalakshmi <sup>1</sup>, P. Shanmugavadivu <sup>1</sup> and S. Vijayalakshmi <sup>2,\*</sup>

<sup>1</sup> Department of Computer Science and Applications, Gandhigram Rural Institute (Deemed to be University), Dindigul, Tamilnadu, India; avijigri@gmail.com (A.V.); psvadivu67@gmail.com (P.S.)

<sup>2</sup> School of Sciences, Christ University, Pune Lavasa Campus, Lavasa, Maharashtra, India.

\* Correspondence: svijisuji@gmail.com or s.vijayalakshmi@christuniversity.in; Tel.: +91-9990353845

<sup>†</sup> Presented at the 12th International Electronic Conference on Sensors and Applications (ECSA-12), 12–14 November 2025; Available online: <https://sciforum.net/event/ECSA-12>.

## Abstract

The health and productivity of dairy cows are critical factors in sustainable livestock management. Along with the rapid rise in intelligence and technology, applying intelligence in livestock management helps in monitoring and provide precise and effective care for the cattle herd. This research designs an intelligent system that can assist the farmers and predict GIR cows' diseases and a support system powered by Artificial Intelligence (AI). The proposed system integrates Internet of Things (IoT) and sensors to track and monitor critical health parameters of the GIR cow, which includes the step count, lying time, rumination time, heart rate, and various environmental factors contributing to the well-being of the cow. The data points that are gathered from the sensors is then processed and analyzed using Machine Learning (ML) algorithms, including Random Forest (RF), Decision Tree (DT), Logistic Regression, K-Neighbors, and Support Vector Machine (SVM), to predict abnormalities including diseases such as lameness, mastitis, heat stress, and digestive problems. The AI techniques used in the system involve complex data processing and pattern recognition to identify early signs of diseases. The RF and DT ML models achieved the highest accuracy (100%), while SVM demonstrated robust performance with 94% accuracy. Integrating real-time monitoring with predictive analytics enables early detection of health issues, allowing timely interventions and improving overall herd management. The proposed system enhances cow welfare and optimizes farm productivity but also has the potential to revolutionize the dairy industry. The complex intelligent system provides a reliable and efficient platform for disease prediction and herd management, and can significantly contribute to the sustainability and profitability of dairy farming, thereby shaping the future of the industry.

**Keywords:** GIR cow; disease prediction; machine learning; dairy cattle health analysis; precision dairy farming; artificial intelligence in agriculture

## 1. Introduction

Effectively managing dairy herds is essential for ensuring both animal welfare and the economic sustainability of dairy farming. The health of dairy cows significantly influences their productivity, reproduction, and longevity. The GIR cow breed, a prominent indigenous breed in India, is chosen as the focus of this research due to its resilience, high milk yield, and adaptability to harsh climatic conditions. These unique characteristics of the GIR cow breed make it an ideal candidate for applying the proposed intelligent system for disease prediction and supporting systems using the potential of Artificial Intelligence (AI). With livestock, health issues can compromise productivity and well-being; the spread of diseases is also unpredictable if the animal with the disease symptoms is not addressed promptly. Traditional methods of monitoring cow health, primarily based on visual inspection and manual observation, are often subjective and can lead to delays in identifying illnesses. Hence, there is an increasing emphasis on developing automated and intelligent systems that can provide timely insights into cow health by integrating sensor-based monitoring and AI techniques.

Internet of Things (IoT) technology—with its ability to capture real-time data from the livestock and Machine Learning (ML)—with its ability to process and mine patterns from the data—advancements have opened new avenues for real-time livestock health monitoring and disease prediction. Studies have demonstrated that monitoring physiological and behavioural parameters such as step count, heart rate, lying time, and rumination activity can provide early indicators of health conditions such as lameness, mastitis, heat stress, and digestive disorders (Borchers et al., 2017; Rutten et al., 2013). Farmers could gain valuable insights and can monitor their cow's health status through wearable sensors and taking timely interventions to prevent disease escalation (Adkin et al., 2009). The application of intelligent techniques should be provided and made available for small-scale farmers and farms to widen the impact of the application of intelligence on everyday life. Educating the impact of the application of intelligent techniques and their impact is a challenge for spreading intelligent techniques in farming and real-time monitoring and data analysis to drive further agricultural innovation.

Diseases are common for livestock; various diseases occur in them, and these can be observed with minor changes in the health parameters of the animals. Lameness, characterised by reduced mobility and increased resting periods, is a major welfare issue in dairy cows, causing significant economic losses due to reduced milk yield and fertility. Studies have shown that monitoring lying time and step count can effectively indicate lameness (Neveux et al., 2006). Similarly, mastitis, an inflammatory udder condition, can be identified by monitoring deviations in heart rate, rumination activity, and physical movement (Cavero et al., 2008). Issues including heat stress, which negatively impacts productivity and reproduction in cows, can be detected by tracking variations in heart rate and lying time (West., 2003).

Several studies have explored the application of sensors and intelligence for monitoring cow health; there is a need for a comprehensive, integrated system that monitors multiple health indicators and leverages ML techniques to predict diseases. This study addresses this gap by developing an intelligent system to monitor a prominent variety of cattle in India—the GIR Cow—for its disease prediction and supporting system using AI, which combines IoT-based sensor networks and ML models to predict common health issues in GIR cows.

## 2. Background

The application of intelligent techniques in various applications has gained momentum. Intelligent techniques and solutions are designed and developed for multiple use cases. The application of intelligent techniques has found various implementation options in precision agriculture, where minute factors are considered for farming. Integrating IoT

Academic Editor(s): Name

Published: date

**Citation:** Vijayalakshmi, A.; Shanmugavadivu, P.; Vijayalakshmi, S. Smart GIR Cow's Disease Prediction & Supporting System Using Artificial Intelligence. *Eng. Proc.* **2025**, volume number, x. <https://doi.org/10.3390/xxxxx>

**Copyright:** © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

and ML in livestock management has gained significant traction. Research and development efforts have increasingly focused on developing systems to monitor the health and behaviour of dairy cows, thereby enhancing productivity, welfare, and farm efficiency. Critical contributions are made in cattle health monitoring and disease prediction using technological interventions by various researchers in various contexts. Farm monitoring and predictive systems are required to control and monitor the changes in the livestock and their maintenance. The intelligent systems enable the farmers to know the status of their herd well in advance and plan the strategies of medication and appropriate changes required in the farm, helping them to make informed decisions. Emerging IoT technologies have been widely applied to monitor cows' behaviour and physiological parameters. The effectiveness of accelerometers and gyroscopes in tracking activities such as walking, lying, and standing (Borchers et al., 2017). The study demonstrated real-time monitoring of activities helps them with the identification of behavioural changes associated with health conditions including lameness and heat stress. Wearable sensor networks could monitor the rumination and lying behaviour, establishing correlations between these parameters; health indicators including the digestive disorders and stress (Rutten et al. 2013). Sensors combined with the AI-ML techniques in gathering data and processing them plays a significant role in planning and implementing intelligent farming techniques.

Different sensors capture different parameters that are associated with the health, environment, and daily routines according to the needs of the farm and livestock managed. The health and other environment parameters of each of the animal in the herd are collected with the help of the intelligent techniques for analyzing and processing. Accurate data is to be collected from the sources (i.e., sensors), and preprocessing techniques prepare and support in validating the data before storing in the system. In order to enhance the accuracy of health monitoring, various studies have combined multiple and redundant sensors to collect a range of data points. Photoplethysmography (PPG) sensors and rumination monitoring devices detect early signs of mastitis in dairy cows (Cavero et al., 2008). Their findings underscored the importance of combining sensors to capture the heart rate variability and rumination patterns to achieve better understanding of cow health. Pressure sensors and Inertial Measurement Units (IMUs) could analyze the cow gait and predict lameness based on deviations from standard mobility patterns (Neveux et al., 2006).

The sensor values, preprocessing, computation, and validation techniques have played an essential role in measuring the livestock's health and other crucial factors that promote sustainable, precise intelligent approaches to farming and agricultural practices. Application of ML techniques has revolutionized the disease prediction in livestock. With the support of Big-Data processing techniques, large volume of sensor-generated data could be easily processed by the ML models, and it promotes to identify complex patterns often been missed by traditional methods. Random Forest (RF), has the ability to handle non-linear relationships between features and the algorithm prevent overfitting through ensemble learning (Borchers et al., 2017). ML and DL approaches are been used by various other researchers for cattle disease classification (Rony et al., 2021; Pavkin et al., 2021). Various other researchers in similar settings have used RF models to predict multiple diseases in dairy cows with high accuracy and interpretability Rutten et al. (2013). The application of ML to the data helps to identify and act upon the hidden data patterns. These support the generation of a more robust and intelligent system and promote the application of intelligent techniques in farming and agricultural practices. AI and its applications can also be implemented in multiple segments to ensure that the entire system uses the maximum potential of the systems. Intelligence that comes from data analysis without intelligent techniques is a challenge, where the machine can effectively scroll through abnormalities in the data and act upon the observations. Support Vector Machines (SVMs)

have also been employed for binary classification tasks, such as distinguishing healthy cows from those suffering from specific ailments.

Cavero et al. (2008) implemented SVM to detect mastitis in dairy cows by analysing heart rate and activity levels. The study demonstrated that SVM could effectively handle high-dimensional data, achieving robust results in differentiating between normal and infected cows. Additionally, Adkin et al. (2009) utilised Logistic Regression to estimate the likelihood of lameness based on movement and rumination data, showing the model's efficiency in handling linear relationships between features. ML algorithms are applied and have been applied for multiple applications on similar lines; a complex system is to be designed to help analyse and process large amounts of data effectively and securely, ensuring the reliability of the data.

Despite the advancements made in health monitoring systems, several challenges persist. One of the primary limitations lies in the accuracy of individual sensors, which can be affected by external factors such as environmental conditions and sensor placement (Rutten et al., 2013). Furthermore, most studies have focused on detecting single diseases in isolation, with limited emphasis on comprehensive systems capable of predicting multiple diseases simultaneously mentioned in papers including (Li et al., 2024, Bhanuteja et al., 2021, Gloster et al., 2011). There is also a need for scalable solutions that can handle large herds without compromising the quality of predictions. The primary objective of this study is to design and implement a system that enables real-time health monitoring and disease prediction in GIR cows. By employing IoT sensors to collect vital health parameters (Džermeikaitė et al., 2023) and utilising ML algorithms (Turimov et al., 2023) such as RF, Decision Tree (DT), Logistic Regression, KNeighbors, and SVM, the proposed system aims to achieve accurate and timely identification of diseases like lameness, mastitis, heat stress, and digestive disorders. Integrating the advancements in technology not only enhances the welfare of the livestock but also improves the efficiency, sustainability and profitability of farming.

The AI-based Intelligent GIR Cow's Disease Prediction & Supporting System presented in the study aims to address the gaps and challenges identified by integrating multiple IoT sensors, with the support of the AI-ML algorithms to monitor live health parameters on the livestock. The proposed solution could easily identify and predict diseases, including lameness, mastitis, heat stress, and digestive problems. Apart from the existing systems, the proposed system works on the basis of real-time data. The system combines various predictions with the support of ML algorithms like RF, DT, SVM, K-NN, and Logistic Regression and provide explainability of the predictions for the end-users. The environment factor monitoring and estrus detection expand the system functionality, enabling farmers to make informed health management and reproductive planning decisions.

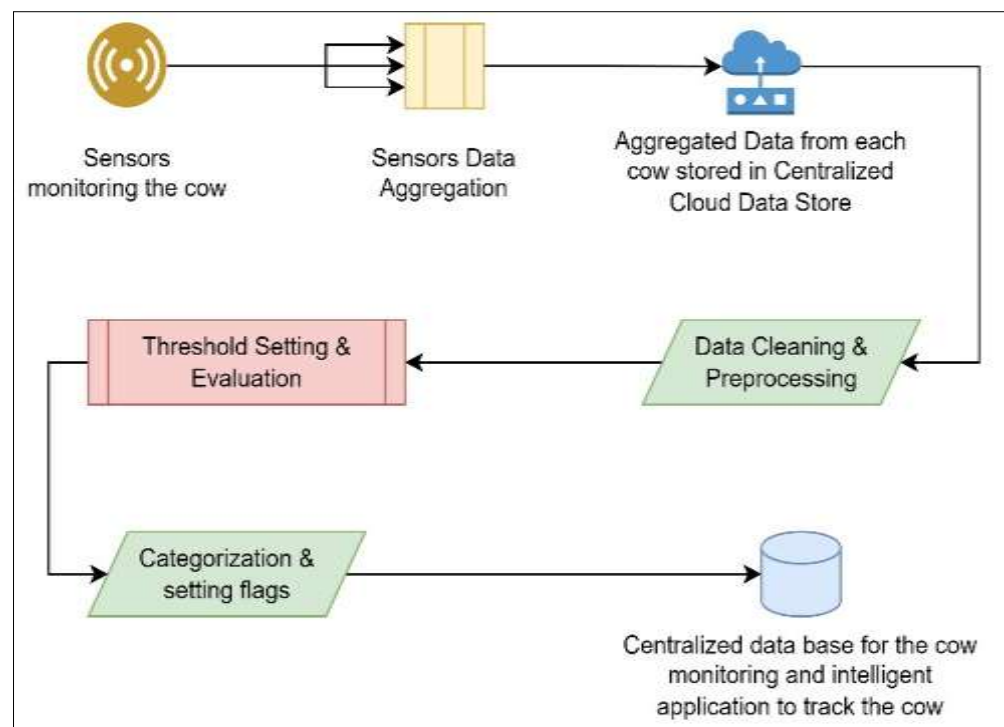
### 3. Proposed Model

The proposed Disease Prediction & Supporting System for the GIR cow, integrates various IoT sensors and the data been processed by automated pipelines and ML algorithms help to achieve real-time monitoring, prediction and support system for GIR cows. The proposed system consists of these major phases:

1. Data collection
2. Data processing
3. Feature extraction
4. ML application
5. Monitoring Dashboard

This section describes each phase in detail, including the sensor setup, data aggregation, and the ML models employed.

The data from the sensors employed in the cow helps gather the data; the data from the sensors are then aggregated and stored in a centralised shared database that can handle data on a large scale. The data is then updated with the appropriate clearing, processing and pre-processing steps to make it convenient for future applications. The data then passes through a threshold setting and evaluation mechanism that has been seen by the experts and veterinarians who assess the cattle and give the required changes in the parameters and thresholds. Based on the mechanisms, threshold flags will be generated to help the farmers and farm managers attend to that cattle and provide the required support in addressing the change observed concerning the parameters. The processed and labelled data is then stored in another database for processing and informed decision-making.



**Figure 1.** The overall methodology of the intelligent system.

### 3.1. Data Collection

The first phase of the methodology of the proposed system to monitor the cattle in the herd involves continuous monitoring of health parameters with the help of the combination of wearable sensors attached to the cows. The data was gathered from Rishi Dairy Farms, a farm of GIR cows which is located in Seelapadi, Dindigul, TamilNadu, India; the sensors were attached to a set of cows for one week and then gathered and stored for the following parameters:

- Cow\_ID
- Date
- Step\_Count
- Distance\_Traveled\_km
- Lying\_Time\_hrs
- Lying\_Standing\_Transitions
- Resting\_Periods\_hrs
- Active\_Periods\_hrs
- Heart\_Rate\_bpm
- Body\_Temperature\_C

- Rumination\_Time\_hrs
- GPS\_Latitude
- GPS\_Longitude
- Estrus\_Activity
- Illness\_Alert
- Body\_Weight\_Estimate\_kg
- BCS
- Health\_Status

The selected sensors and their placements are determined based on the specific parameters required to predict diseases. The critical sensors used in this study include:

- 3-axis Accelerometer: Captures data on step count and distance travelled. It is strategically placed on the cow's leg or neck to monitor mobility.
- Gyroscope: Tracks lying/standing transitions to detect changes in posture.
- IMU (Inertial Measurement Unit): Combines accelerometer and gyroscope data to analyse gait patterns and identify mobility issues.
- PPG/ECG Sensor: Measures the cow's heart rate to detect variations that may indicate illness or stress.
- Microphone + Accelerometer: Used to monitor rumination activity, placed near the cow's jawline to capture chewing behaviour.
- Temperature and Humidity Sensors: Monitors ambient conditions to detect potential heat stress

Table 1 explains the different types of sensors and their placement with the justification for the parameters that can be monitored from the placement of sensors. Each cow has these sensors, continuously transmitting data to a centralised cloud storage system. The sensors are connected via a wireless communication protocol, ensuring reliable data transmission even in remote farm settings.

**Table 1.** Sensor(s) and the parameters and the expected placement of the sensors.

| Sensor(s)                       | Parameter                      | Placement                |
|---------------------------------|--------------------------------|--------------------------|
| 3-axis Accelerometer            | Step Count, Distance Travelled | Leg/Neck Pedometer       |
| Gyroscope                       | Lying/Standing Transitions     | Leg/Neck                 |
| IMU (Accelerometer + Gyroscope) | Resting/Active Periods         | Leg or Collar            |
| IMU + Pressure Sensors          | Gait Analysis                  | Leg/Foot                 |
| Thermistor/Infrared Sensor      | Temperature Monitoring         | Ear Tag or Rectal Sensor |
| Microphone + Accelerometer      | Rumination Monitoring          | Near Jawline             |
| GPS Module                      | GPS Tracking                   | Collar or Ear Tag        |
| PPG/ECG Sensor                  | Heart Rate Monitoring          | Chest or Neck Strap      |
| DHT22 + Light Sensor            | Environmental Monitoring       | Barn or Collar           |
| Accelerometer + Thermistor      | Estrus Detection               | Leg + Vaginal Sensor     |

### 3.2. Data Aggregation & Preprocessing

The data collected from the IoT sensors is aggregated and stored in a centralised database. The aggregation process consolidates the real-time data streams from each cow and organises them for further analysis. The key steps involved in this phase include:

1. Data Cleaning: Removing erroneous or noisy data points caused by sensor malfunctions or environmental interference. For example, the extreme values in the step count or heart rate that exceed physiological limits are flagged and removed.

2. Data Imputation: Missing data with respect to some of the system errors can be handled by employing data interpolation techniques, ensuring that the dataset remains consistent for models.
3. Feature Extraction: Extracting the relevant features from the raw sensor data. These features include the calculation of average step count, lying time, rumination time, number of lying/standing transitions, and heart rate variability.

The Body Condition Scoring (BCS) is crucial part of the system that create and evaluate the cow's overall health and nutritional parameters. The BCS is measured on a scale of 1 to 5, where 1 indicates a weak cow, and 5 denotes an over-conditioned cow. The scoring is based on visual inspection, along with the feature extraction based on AI techniques and physical observation of anatomical regions to assess fat and muscle reserves. Comprehensive BCS is calculated using sensor-based parameters and the physical inspection; the data points including step count, lying time, rumination time, and heart rate etc. also contribute the BCS calculation.

$$BCS = f(\text{Step Count, Lying Time, Rumination Time, Heart Rate}) \quad (1)$$

Function  $f$  is based on the following approach, with weights assigned for each of the parameters, i.e., for step count, the weight is 20%; for lying time, it is 30%; for rumination time, it is 35%; and for heart rate, it is 15%.

- Low Activity (Step Count < 4,000) + High Lying Time (> 12 h) + Low Rumination Time (< 7 h) → Lower BCS Score.
- Moderate Activity (Step Count 4,000 to 7,000) + Optimal Lying Time (10-12 h) + Sufficient Rumination Time (7-9 h) → Ideal BCS Score.
- High Activity (Step Count > 7,000) + Low Rumination Time (< 6 h) or High Lying Time (> 12 h) → higher BCS or potentially over-conditioned Score.

### 3.3. Threshold Setting and Illness Detection

Based on the extracted features, the system employs threshold-based rules to detect abnormal behaviour patterns. These rules are derived from existing literature on cow health monitoring and expert consultations. The thresholds set for critical parameters are as follows:

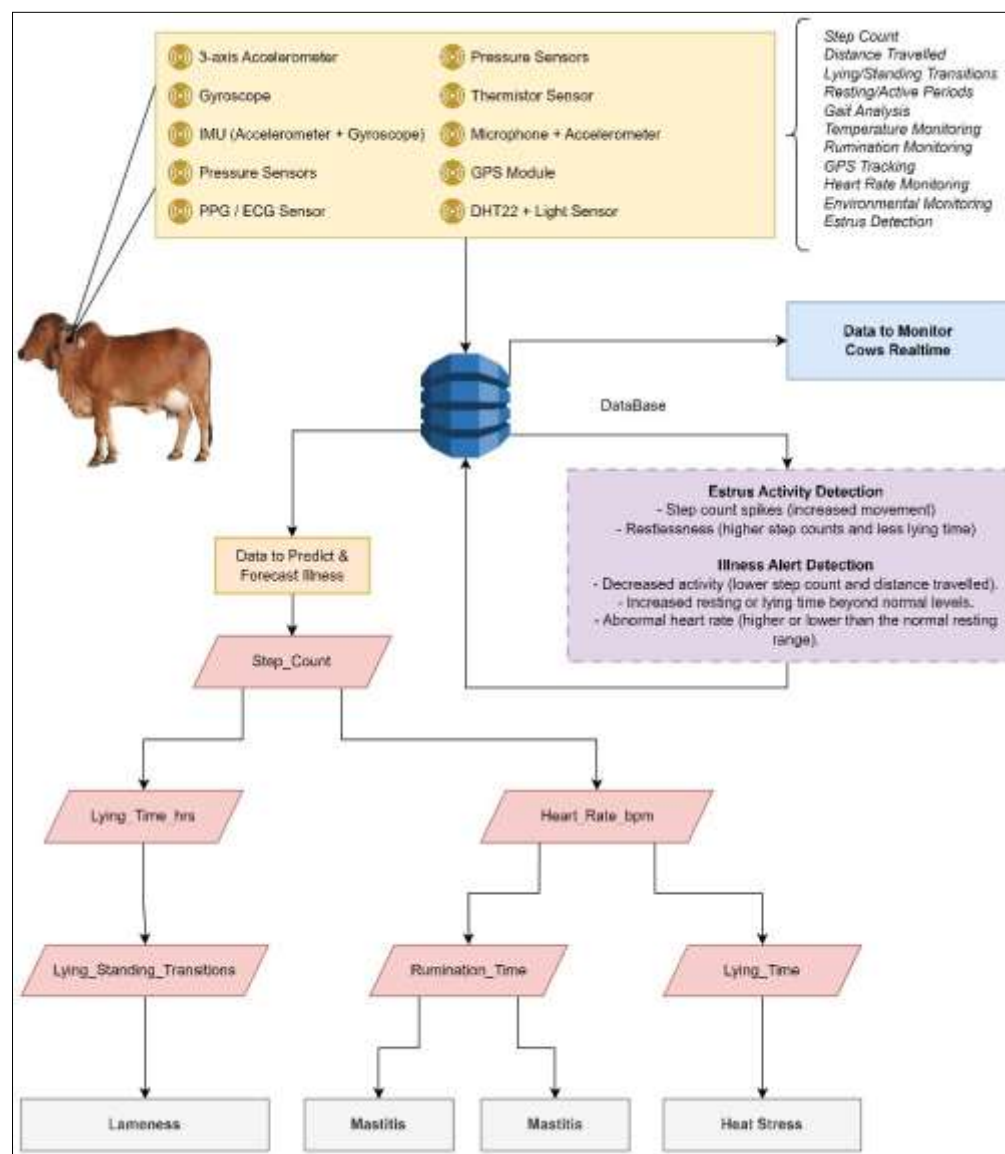
- Step Count below 4,000 steps/day and lying time over 12 h indicate lameness.
- Heart Rate exceeding 75 bpm with rumination time below 7 h/day suggests mastitis.
- Heart Rate above 80 bpm and lying time over 12 h signal heat stress.
- Rumination Time below 5 h/day combined with a step count below 3,800 steps/day points to digestive issues.

These thresholds allow the system to flag potential illnesses and categorise cows based on their predicted health status. Various diseases and the parameters considered to determine whether a cattle can potentially have been prone to any of the following disease flags are based on monitoring the change in the following parameters.

- **Lameness:** Low step count, high lying time, and few transitions.
- **Mastitis:** Elevated heart rate, low rumination time, and reduced activity.
- **Heat Stress:** High heart rate, increased lying time, and reduced activity.
- **Digestive Problems:** Reduced rumination time, elevated heart rate, and low step count.
- **Healthy:** Cows that do not match disease conditions are classified as healthy.

Figure 2 shows the detailed methodology of the proposed intelligent system to monitor cattle herds in farms. The diagram also shows the alert detection system and centralised data store that captures the data points for intelligent analysis and predictive

modelling. The preprocessing system also allows the system to flag the cattle into groups. It suggests that managers should take special care of cattle with significant changes in the monitored parameters. Different illness flags are also assigned to cattle if they fall into certain groups. This intelligent system helps to isolate the cattle with issues and gives unique, precise treatments according to their requirements to get them to normal. The real-time monitoring and intelligent systems and their application help minimise the health risks and diseases that spread in the herd.



**Figure 2.** Detailed methodology and flag generation on illness.

Table 2 shows the sample dataset generated for the study. Data from each cow for one week is recorded, and some attributes are determined based on the sensor readings and processing of the data from the sensors.

**Table 2.** Dataset Used for the study.

| Step_Count | Distance_Traveled_km | Lying_Time_hrs | Lying_Standing_Transitions | Resting_Periods_hrs | Active_Periods_hrs | Heart_Rate_bpm |
|------------|----------------------|----------------|----------------------------|---------------------|--------------------|----------------|
| 4981       | 2.63                 | 11.01          | 8                          | 9.03                | 14.97              | 80.62          |
| 4178       | 3.64                 | 11.75          | 11                         | 8.5                 | 15.5               | 69.45          |

|      |      |       |    |       |       |       |
|------|------|-------|----|-------|-------|-------|
| 4779 | 3.21 | 13.33 | 7  | 7.62  | 16.38 | 78.77 |
| 4687 | 2.97 | 14.2  | 9  | 8.97  | 15.03 | 80.83 |
| 4100 | 3.89 | 11.85 | 9  | 10.15 | 13.85 | 80.1  |
| 4576 | 2.65 | 10.47 | 11 | 7.43  | 16.57 | 65.27 |
| 4061 | 2.5  | 12.61 | 11 | 10.38 | 13.62 | 84.17 |
| 2863 | 2.71 | 10.18 | 8  | 9.83  | 14.17 | 72.92 |
| 3607 | 3.3  | 13.5  | 11 | 8.92  | 15.08 | 64.99 |

| Body_Tem<br>perature_C | Rumination<br>_Time_hrs | Estrus_<br>Activity | Illness_Alert | Body_Weight_<br>Estimate_kg | BCS | Health_<br>Status |
|------------------------|-------------------------|---------------------|---------------|-----------------------------|-----|-------------------|
| 38.42                  | 7.85                    | FALSE               | FALSE         | 405.26                      | 2.5 | Healthy           |
| 37.82                  | 6.99                    | FALSE               | FALSE         | 407.28                      | 3   | Healthy           |
| 39.2                   | 8.34                    | FALSE               | FALSE         | 406.42                      | 4   | Healthy           |
| 38.16                  | 5.68                    | FALSE               | FALSE         | 405.94                      | 2.5 | Healthy           |
| 38.1                   | 7.62                    | TRUE                | FALSE         | 407.78                      | 2.5 | Healthy           |
| 39.68                  | 8.98                    | FALSE               | FALSE         | 405.3                       | 3   | Healthy           |
| 38.47                  | 9.67                    | FALSE               | FALSE         | 405                         | 4   | Heat Stress       |
| 38.37                  | 7.67                    | TRUE                | FALSE         | 405.42                      | 3   | Healthy           |
| 38.35                  | 6.98                    | FALSE               | FALSE         | 406.6                       | 3   | Healthy           |

### 3.3. Disease Prediction Using Machine Learning

The final stage of the proposed methodology focuses on developing and training ML models to predict diseases based on the aggregated data. Five ML algorithms were employed to classify the cows into health categories:

**Random Forest (RF):** The RF algorithm is an ensemble learning method that constructs multiple DTs and aggregates their predictions to enhance accuracy. Each decision tree in the forest is trained on a randomly selected subset of the dataset, which helps capture diverse aspects of the data and reduce overfitting. The RF model was applied to the dataset on the features from the dataset including step count, heart rate, lying time, rumination time, and temperature. The model combined the predictions from all DTs using a majority voting mechanism, providing a final output for each cow's health status (Kukreja et al., 2024). The high accuracy of the RF model in detecting diseases such as lameness and mastitis indicates its robustness in handling complex relationships and non-linear patterns in the dataset.

**Decision Tree (DT):** A simple and powerful ML model that splits the dataset into branches based on the conditional rules derived from the features. Each internal node of the tree represented a specific health parameter (for example: heart rate), and each branch corresponded to a decision rule. The leaf nodes signified the predicted disease class (lameness, heat stress). DT could provide interpretable results, allowing users to understand the conditions under which each disease was predicted (Lowie et al., 2021). DT could effectively capture the relationships between parameters such as lying time, step count, and rumination.

**Logistic Regression (LR):** LR is a statistical model commonly used for binary and multi-class classification problems. LR estimates the probability of each of the disease outcome based on a linear combination of the input features. LR could predict the likelihood of a cow being healthy or having a specific disease based on parameters such as step count, rumination time, and lying-standing transitions (Megahed et al., 2022). LR effectively indicates the relationship between specific features and disease conditions. The strength of LR lies in its interpretability and ability to provide probabilistic insights into the health status of each cow.

**K-Neighbors (K-NN):** K-NN is a non-parametric algorithm that could classify the data points based on the majority class of their  $k$  nearest neighbours. K-NN could classify the cows into different health categories by comparing their feature values with those of their closest neighbours (Garcia et al., 2024; martin et al., 2015). K-NN calculates the Euclidean distance between data points and determine the nearest neighbours, the majority class among these neighbours is then assigned to the cow. K-NN model performed well classifying diseases such as lameness and digestive problems. The ability to detect local patterns in the data made K-NN a practical choice for identifying health conditions by specific combinations of parameters.

**Support Vector Machine (SVM):** SVM, a supervised learning ML algorithm aims to find the optimal hyperplane that separates different classes in the feature space. SVM works by differentiating between healthy cows with the ones who are suffering from diseases based on the input features (Miekley et al., 2013; Mammadova et al., 2013). By maximising the margin between classes, SVM effectively separate normal and sick cows in highdimensional spaces. With the help of non-linear kernel functions SVM handle complex relationships between the features and effectively distinguishes between overlapping classes.

The dataset used in the study, consists of health parameters and disease labels, was divided into training and testing sets. Each model was trained on the training set and evaluated on the testing set. The model performance was assessed using accuracy, precision, recall, and F1-score metrics. The RF and DT models achieved the highest accuracy of 100%, while the SVM model demonstrated robust classification performance with an accuracy of 94%.

The final system integrates the trained machine learning models with a real-time monitoring platform. The system continuously monitors the sensor data, processes it through ML models, and generates alerts when a cow's health deviates from the established norms. The alerts are communicated to farmers and veterinarians via a mobile application, enabling timely interventions and reducing the risk of disease escalation.

## 4. Results and Discussion

The performance of the proposed intelligent system for GIR cow's disease prediction and supporting system was evaluated based on its ability to predict common health issues in GIR cows. The system's predictions were compared with actual disease conditions, and the results are discussed below in terms of model performance, disease classification accuracy, and implications for dairy farming.

The study employed five machine learning algorithms—RF, DT, LR, K-NN, and SVM—to predict diseases such as lameness, mastitis, heat stress, and digestive problems. The accuracy of each of the model was calculated using a test dataset consisting of labelled examples. The evaluation metrics included accuracy, precision, recall, and F1-score to ensure a comprehensive assessment of each model's ability to classify health conditions accurately.

RF and DT models achieved an accuracy of 100%, demonstrating robustness in identifying and learn from the complex patterns within the dataset. These models could successfully capture non-linear relationships and interactions between multiple health parameters of the cow, which includes the step count, lying time, heart rate, and rumination activity. The high accuracy of these models is attributed to their ability to perform feature selection and ensemble learning, which reduced the risk of overfitting and improved generalizability.

SVM attained an accuracy of 94%, and the effectiveness of SVM lies in distinguishing between health conditions is due to the model's capacity to find optimal hyperplanes that separate classes in highdimensional feature spaces. The slightly lower accuracy than RF

and DT models suggests that SVM face limitations in capturing subtle relationships between parameters.

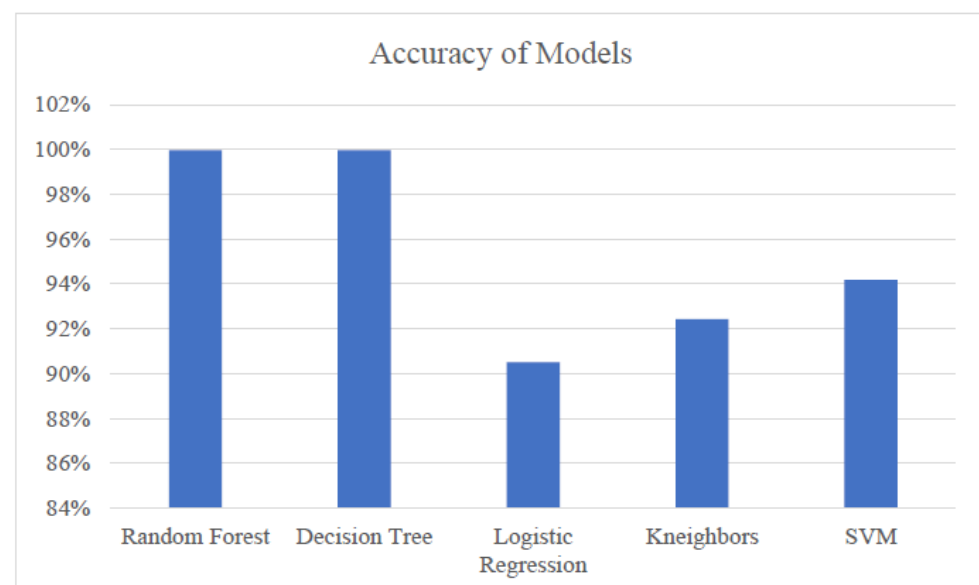
K-NN achieved an accuracy of 92%, indicating that it effectively classified diseases based on feature similarities. The K-NN model was powerful in identifying lameness, which involves straightforward patterns in step count and lying time. However, its reliance on local neighbours may have limited its performance for diseases with overlapping feature distributions.

LR reached an accuracy of 91%, performing well in cases where the relationship between features and health outcomes was approximately linear. However, this model's relative simplicity limited its ability to handle more complex interactions between the parameters.

The results of applying the ML algorithm to predict the health issues from the gathered data for the Gir Cow were done. The data for the various parameters are computed from the sensors' readings and stored in the database for further processing and notification systems. The ML algorithms also use these data to predict diseases and the chances of getting sick and determine whether the cow needs personal attention.

**Table 3.** ML algorithms and their accuracy performance.

| Sl. No. | Models              | Accuracy |
|---------|---------------------|----------|
| 1       | Random Forest       | 100%     |
| 2       | Decision Tree       | 100%     |
| 3       | Logistic Regression | 91%      |
| 4       | K-Neighbors         | 92%      |
| 5       | SVM                 | 94%      |



**Figure 3.** Comparison of Accuracy of the ML models.

Based on established thresholds for each health parameter, the system identified critical diseases affecting GIR cows, including lameness, mastitis, heat stress, and digestive problems. The threshold-based rules were designed to detect abnormalities in step count, heart rate, lying time, and rumination time, which indicate different health conditions. The results demonstrated that:

**Lameness Detection:** The system accurately flagged cows with low step count (<4000 steps/day), high lying time (>12 h/day), and few lying/standing transitions (<5 transitions)

as potentially lame. This finding aligns with previous studies emphasising the importance of monitoring cow mobility and posture changes to identify lameness (Neveux et al., 2006).

**Mastitis Prediction:** Cows exhibiting elevated heart rates ( $>75$  bpm), reduced rumination time ( $<7$  h/day), and decreased activity levels were classified as potentially suffering from mastitis. This classification approach is in line with the findings of Cavero et al. (2008), who also have highlighted the role of physiological indicators in early mastitis detection.

**Heat Stress Identification:** Cows with high heart rates ( $>80$  bpm) and increased lying time ( $>12$  h/day), combined with reduced step count, are flagged to be experiencing heat stress. These results also align with the study by West (2003), who observed that cows who are exposed to high temperatures shows prolonged lying periods and elevated heart rates.

**Digestive Problems:** The system could identify cows with reduced rumination time ( $<5$  h/day), elevated heart rates ( $>70$  bpm), and low step count as cows with digestive issues. This confirms that the significance of rumination monitoring in detecting gastrointestinal disorders in cows.

The results of this study demonstrate that the IoT-based health monitoring system with AI techniques offers a reliable and scalable solution to monitor and get disease alerts. The system could accurately identify and predict common diseases in cows, the proposed system enables the farmers to take preventive steps and timely interventions on sick cows. This approach improves the welfare and enhances milk productivity by minimising the impact of health issues on the herd. The high accuracy achieved by the ML algorithms RF and DT models underscores the effectiveness of ensemble learning techniques. These models can handle complex datasets and make robust predictions based on various health parameters. The combination of threshold-based rules and ML ensures that the system remains interpretable for farmers and veterinarians, allowing them to understand the basis of each prediction and make informed decisions.

## 5. Conclusions

This study designs and presents a comprehensive and complex data-based approach on health monitoring and disease prediction in GIR cows. The system is designed based on the IoT-based sensor networks and ML algorithms. The system integrates various sensors to continuously monitor key health parameters including step count, lying time, heart rate, rumination time, and temperature. With the support of real-time data, the system employs various ML models—RF, DT, LR, K-NN, and SVM—to predict diseases including lameness, mastitis, heat stress, and digestive problems. The findings from the experiments made as part of the system demonstrate that RF and DT models achieved accuracy of 100%, while SVM demonstrated robust classification with 94% accuracy. The study confirms that integrating ensemble learning techniques with IoT-based health monitoring provides reliable predictions, enabling early disease detection and timely interventions. The threshold-based classification rules, with ML models makes the system both accurate and interpretable. The system offers significant benefits on dairy farming, enhanced cattle welfare, improved milk productivity, and reduced economic losses. The combination of data-driven insights and real-time monitoring enables a proactive approach to livestock management, contributing to the sustainability and profitability of dairy operations. The future work will focus on expanding the system capabilities to include more physiological and behavioral parameters for more precise predictions. Integrating Deep Learning (DL) techniques with explainability will enhance the system ability to capture complex hidden patterns in behavior and health. Further developments could also make the system

scalable and cost-effective for larger herds, benefiting a more comprehensive range of dairy farms.

**Author Contributions:**

**Funding:**

**Institutional Review Board Statement:**

**Informed Consent Statement:**

**Data Availability Statement:**

**Conflicts of Interest:**

## References

1. Adkin, A.; Green, M.; Hudson, C.; Bradley, A. Monitoring lameness in dairy cattle: The relationship between environmental factors and lameness. *J. Dairy Sci.* **2009**, *92*, 1802–1808.
2. Borchers, M.R.; Chang, Y.M.; Tsai, I.C.; Wadsworth, B.A.; Bewley, J.M. Machine learning in dairy cattle behavior and health monitoring. *J. Dairy Sci.* **2017**, *100*, 9805–9823.
3. Cavero, D.; Tölle, K.H.; Buxadé, C.; Krieter, J. Mastitis detection in dairy cows using support vector machines. *Livestig. Sci.* **2008**, *114*, 280–286.
4. Neveux, S.; Weary, D.M.; Rushen, J.; Von Keyserlingk, M.A.G.; de Passillé, A.M. Hoof discomfort and its effects on resting behavior in dairy cows. *Appl. Anim. Behav. Sci.* **2006**, *101*, 120–127.
5. Rutten, C.J.; Velthuis, A.G.; Steeneveld, W.; Hogeveen, H. Invited review: Sensors to support health management on dairy farms. *J. Dairy Sci.* **2013**, *96*, 1928–1952.
6. West, J.W. Effects of heat-stress on production in dairy cattle. *J. Dairy Sci.* **2003**, *86*, 21312144.
7. Li, D.; Li, B.; Li, Q.; Wang, Y.; Yang, M.; Han, M. Cattle identification based on multiple feature decision layer fusion. *Sci Rep* **2024**, *14*, 26631. <https://doi.org/10.1038/s41598-024-76718-x>.
8. Bhanuteja, T.; Kumar, K.V.N.; Poornachand, K.S.; Ashish, C.; Anudeep, P. Symptoms Based Multiple Disease Prediction Model using Machine Learning Approach. *Int. J. Innov. Technol. Explor. Eng. (IJITEE)* **2021**, *10*, 2278–3075.
9. Gloster, J.; Ebert, K.; Gubbins, S.; Bashiruddin, J.; Paton, D.J. Normal variation in thermal radiated temperature in cattle: Implications for foot-and-mouth disease detection. *BMC Vet. Res.* **2011**, *7*, 1–10.
10. Rony, M.; Barai, D.; Hasan, Z. Cattle external disease classification using deep learning techniques. In Proceedings of the 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 6–8 July 2021; pp. 1–7.
11. Pavkin, D.Y.; Dorokhov, A.S.; Vladimirov, F.E.; Dovlatov, I.M.; Lyalin, K.S. Algorithms for detecting cattle diseases at early stages and for making diagnoses and related recommendations. *Appl. Sci.* **2021**, *11*, 11148.
12. Džermeikaitė, K.; Bačėninaitė, D.; Antanaitis, R. Innovations in cattle farming: Application of innovative technologies and sensors in the diagnosis of diseases. *Animals* **2023**, *13*, 780.
13. Turimov Mustapoevich, D.; Muhamediyeva Tulkunovna, D.; Safarova Ulmasovna, L.; Primova, H.; Kim, W. Improved Cattle Disease Diagnosis Based on Fuzzy Logic Algorithms. *Sensors* **2023**, *23*, 2107.
14. Kukreja, V.; Srivastava, P.; Garg, A.; Hariharan, S. Synergizing CNN and Random Forest for Accurate Cattle Disease Identification. In Proceedings of the 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), Bangalore, India, 28–29 June 2024; pp. 1–6.
15. Lowie, T.; Callens, J.; Maris, J.; Ribbens, S.; Pardon, B. Decision tree analysis for pathogen identification based on circumstantial factors in outbreaks of bovine respiratory disease in calves. *Prev. Vet. Med.* **2021**, *196*, 105469.
16. Megahed, A.; Kandeel, S.; Alshaya, D.S.; Attia, K.A.; AlKahtani, M.D.; Albohairy, F.M.; Selim, A. A comparison of logistic regression and classification tree to assess brucellosis associated risk factors in dairy cattle. *Prev. Vet. Med.* **2022**, *203*, 105664.
17. García, R.; Aguilar, J. A meta-learning approach in a cattle weight identification system for anomaly detection. *Comput. Electron. Agric.* **2024**, *217*, 108572.
18. Martin, M.K.; Helm, J.; Patyk, K.A. An approach for de-identification of point locations of livestock premises for further use in disease spread modeling. *Prev. Vet. Med.* **2015**, *120*, 131–140.

19. Miekley, B.; Traulsen, I.; Krieter, J. Mastitis detection in dairy cows: The application of support vector machines. *J. Agric. Sci.* **2013**, *151*, 889–897.
20. Mammadova, N.; Keskin, İ. Application of the support vector machine to predict subclinical mastitis in dairy cattle. *Sci. World J.* **2013**, *2013*, 603897.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.