International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Intelligent Prediction Of Environmental Stress And Disease Risk In Sustainable Livestock Systems

Baliram N. Gaikwad[1], Arjit Tomar[2], Swati Nikam[3], Santosh Chobe[4], Dr. Jayant Damodar Supe[5], Amruta Mhatre[6], Pratik Mungekar[7]

[1]Department of Lifelong Learning and Extension, University Of Mumbai, Maharashtra, India. gaikwadbn@gmail.com
[2]Department of Computer Science & Engineering,Noida international University, Greater Noida, Uttar Pradesh 203201, India,. arjit.tomar@niu.edu.in
[3]Department of Computer Engineering, Pimpri Chinchwad College of Engineering and Research, Pune, Maharashtra, India. swatinikam3@gmail.com
[4]Department of Information Technology, Pimpri Chinchwad College of Engineering and Research, Pune, Maharashtra, India. sanchobe@gmail.com
[5]Department of Computer Science and Engineering, Rungta College of Engineering and Technology, Rungta International Skills University, Bhilai, Chhattisgarh – 490024, India. dr.jayant.supe@rungta.org
[6]Department of Computer Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, Maharashtra, India. amruta.mhatre@bvcoenm.edu.in
[7]Director, Research Innovation and Internationalization, Sharda Education Society Thane, University of Mumbai, Mumbai, Maharashtra, India. pratikmngkr11@gmail.com

Abstract
Environmental stress and disease occurrence are major challenges affecting animal welfare, productivity, and sustainability in modern livestock production systems. This study proposes an intelligent machine learning-based framework for predicting environmental stress and associated disease risk through the integrated analysis of environmental, behavioral, physiological, and health-related livestock data. The framework incorporates data acquisition, preprocessing, feature engineering, environmental stress prediction, disease-risk assessment, and sustainability-oriented decision support. Environmental variables such as temperature, humidity, air quality, ventilation conditions, and Temperature-Humidity Index are analyzed together with animal-level indicators including body temperature, respiration rate, activity, feeding behavior, water consumption, and rumination. Multiple machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, Gradient Boosting, and Artificial Neural Network, are comparatively evaluated. The results indicate that Gradient Boosting provides the strongest predictive performance, achieving 95.21% accuracy for environmental stress prediction and 94.36% accuracy for disease-risk prediction. The proposed framework enables early identification of potentially unfavorable conditions and supports preventive interventions related to environmental control, animal monitoring, and health management. The integration of intelligent prediction with livestock monitoring can contribute to improved animal welfare, reduced health-related losses, efficient resource utilization, and more sustainable livestock production.

Keywords: Sustainable Livestock Systems, Environmental Stress, Disease Risk Prediction, Machine Learning, Precision Livestock Farming, Animal Welfare, Intelligent Decision Support.

1. Introduction

Livestock production is an essential component of the global agricultural system, contributing significantly to food security, rural livelihoods, employment, and economic development. The increasing demand for animal-derived products has encouraged the development of intensive livestock production systems capable of improving productivity and resource utilization. However, livestock productivity and animal health are strongly influenced by environmental conditions, including temperature, relative humidity, air quality, ventilation, water availability, housing conditions, and seasonal variations. Prolonged exposure to unfavorable environmental conditions can induce physiological stress, reduce feed intake, affect reproductive performance, weaken immune responses, and increase susceptibility to diseases [1]. Therefore, continuous assessment of environmental stress and associated disease risks is important for maintaining animal welfare and achieving sustainable livestock production.

1.1 Background and Motivation

Environmental stress represents a major challenge in modern livestock management, particularly under changing climatic conditions. Heat stress is among the most prominent environmental concerns because elevated temperature and humidity can disrupt the thermoregulatory mechanisms of animals. In addition to thermal conditions, poor air quality, excessive concentrations of harmful gases, inadequate ventilation, overcrowding, and unsuitable housing conditions can negatively influence animal health [2]. These environmental factors often interact with animal-specific characteristics such as age, breed, body condition, activity, feeding behavior, and physiological status, making livestock health assessment a complex and multidimensional problem.

Conventional livestock management primarily depends on periodic physical inspection and the experience of farmers or veterinary professionals. Although these approaches remain important, they may not provide sufficiently early indications of environmental stress or emerging disease conditions. Symptoms may become clearly observable only after the animal has experienced substantial physiological disturbance. The increasing availability of Internet of Things (IoT) devices, environmental sensors, wearable technologies, automated monitoring systems, and livestock management databases provides opportunities to continuously collect environmental, behavioral, and physiological information [3]. Such data can support intelligent prediction models capable of recognizing patterns associated with stress and disease before severe consequences occur.

1.2 Environmental Stress and Livestock Health

Livestock animals continuously interact with their surrounding environment, and changes in environmental conditions can directly influence their physiological and behavioral responses. Temperature, humidity, air quality, solar radiation, rainfall, and housing conditions are among the major environmental variables associated with livestock stress. For example, high temperature combined with elevated relative humidity may increase thermal load and result in heat stress. Animals experiencing thermal stress may exhibit increased respiration, reduced feed consumption, altered activity, decreased milk or meat production, and impaired reproductive performance [4].

Environmental stress can also contribute indirectly to disease development by weakening immune responses and creating conditions favorable for pathogens. Poor ventilation and elevated concentrations of ammonia or other gases may increase respiratory problems, while excessive humidity can support microbial growth within livestock housing environments. Consequently, environmental stress and disease risk should not always be considered as independent management problems [5]. Their combined analysis can provide a more comprehensive representation of livestock health and enable preventive intervention before conditions become critical.

1.3 Role of Intelligent Prediction in Sustainable Livestock Systems

Advances in artificial intelligence (AI) and machine learning (ML) have created opportunities to transform livestock monitoring from reactive observation toward predictive and preventive management. Intelligent prediction models can process large volumes of heterogeneous data collected from environmental sensors, wearable devices, animal records, and farm management systems [6]. By identifying complex relationships among environmental conditions, physiological indicators, behavioral patterns, and historical disease occurrences, these models can estimate the probability of environmental stress and potential disease risk. Machine learning techniques such as logistic regression, decision trees, random forests, support vector machines, gradient boosting, and artificial neural networks can be applied to livestock monitoring data for classification and prediction. More advanced deep learning approaches can further capture nonlinear and temporal relationships in continuously collected sensor data. The predicted information can subsequently be transformed into risk levels, alerts, and management recommendations for farmers and livestock managers.

The integration of intelligent prediction with sustainable livestock management is particularly important because early intervention can reduce unnecessary resource consumption, prevent productivity losses, improve animal welfare, and support efficient farm operations. Predictive information may assist farmers in adjusting ventilation, cooling, feeding, water supply, housing conditions, and veterinary interventions according to identified risks [7]. Thus, intelligent prediction can contribute simultaneously to animal health, farm productivity, resource efficiency, and environmental sustainability.

1.4 Research Problem and Objectives

Despite significant developments in precision livestock farming, environmental monitoring and animal health prediction are frequently addressed as separate analytical tasks. Environmental monitoring systems commonly focus on measuring climatic or housing parameters, whereas disease prediction approaches

often concentrate primarily on clinical or physiological information. Such isolated analysis may overlook the relationship between environmental stress and subsequent health deterioration.
The present study therefore focuses on developing an intelligent data-driven framework for the integrated prediction of environmental stress and disease risk in sustainable livestock systems. The framework considers environmental, animal-related, behavioral, and health indicators to generate predictive information that can support timely livestock management decisions. The major objectives are to identify important environmental and animal-related risk factors, preprocess and integrate heterogeneous livestock monitoring data, develop machine learning-based models for stress and disease-risk prediction, evaluate their predictive performance, and translate model outputs into actionable decision-support information.

1.5 Contributions of the Study

The principal contribution of this study is an integrated intelligent prediction framework that establishes a relationship between environmental conditions, livestock stress, and potential disease risk. Unlike monitoring approaches that primarily report current sensor measurements, the proposed approach emphasizes predictive assessment to identify potentially unfavorable conditions before substantial effects on animal health and productivity occur. The framework integrates environmental and livestock-related information through systematic preprocessing and feature engineering and applies machine learning techniques for environmental stress and disease-risk estimation.
Furthermore, the study evaluates the predictive models using appropriate performance measures and comparative analysis to determine their effectiveness and reliability. The resulting predictions are incorporated into a sustainability-oriented decision-support mechanism that can provide risk assessment and management recommendations. Through the combined consideration of animal welfare, disease prevention, environmental conditions, and resource-efficient management, the proposed approach aims to support the transition toward more intelligent, preventive, and sustainable livestock production systems.

2. Literature Review

The increasing adoption of digital technologies in livestock farming has encouraged research on environmental monitoring, animal health assessment, disease prediction, and intelligent farm management. Existing studies have investigated environmental stress indicators, sensor-based monitoring, machine learning models, and precision livestock farming as mechanisms for improving animal health and productivity. However, environmental stress prediction and disease-risk assessment are frequently addressed independently. This section reviews the major developments relevant to the proposed intelligent prediction framework and identifies the research gaps associated with integrating environmental, behavioral, and health information for sustainable livestock management.

2.1 Environmental Stress Factors in Livestock Systems

Environmental conditions have a substantial influence on livestock health, productivity, and welfare. Temperature and relative humidity are among the most frequently investigated variables because their combined effects determine the thermal load experienced by animals. The Temperature-Humidity Index (THI) is widely used as an indicator for assessing heat-stress conditions in cattle and other livestock species. Increasing THI levels have been associated with changes in respiration, feed consumption, milk production, reproductive efficiency, and general animal behaviour [8].
Environmental stress is nevertheless influenced by more than temperature and humidity. Ventilation, air velocity, solar radiation, stocking density, water availability, dust, and concentrations of gases such as ammonia and carbon dioxide can affect the microenvironment surrounding livestock. Poorly controlled housing environments may expose animals to prolonged thermal or respiratory stress [9]. Seasonal changes can further modify these conditions and produce considerable variations in animal responses.
Recent livestock management approaches therefore increasingly rely on continuous environmental measurements rather than occasional observations. Sensor networks can capture changes in temperature, humidity, air quality, and other environmental variables at short intervals, providing a more detailed representation of livestock exposure. However, monitoring environmental parameters alone does not necessarily indicate how individual animals respond to these conditions. This limitation motivates the integration of environmental measurements with behavioral and physiological indicators.

2.2 Disease Risk and Animal Health Monitoring

Disease prevention is a major requirement for sustainable livestock production because disease outbreaks can reduce productivity, increase veterinary costs, affect animal welfare, and generate substantial economic losses. Conventional health assessment typically relies on physical examination, visual symptoms, laboratory tests, and veterinary diagnosis [10]. Although these methods are essential for confirming diseases, their effectiveness for continuous early-risk detection may be limited.

Research in precision livestock farming has therefore examined measurable indicators that can provide early evidence of health deterioration. These include body temperature, heart rate, respiration rate, feeding frequency, water consumption, rumination, movement, posture, sleeping behavior, and changes in production. Abnormal variations in these indicators may occur before clearly visible clinical symptoms develop [11].

Disease risk may also be influenced by environmental conditions. Thermal stress can reduce immune efficiency, while poor ventilation and high concentrations of airborne contaminants may increase susceptibility to respiratory disorders [12]. High humidity and unsuitable housing conditions can facilitate microbial growth and pathogen transmission. Consequently, combining environmental and animal-level indicators may improve the identification of situations associated with increased disease risk.

2.3 IoT and Sensor-Based Livestock Monitoring

The Internet of Things has become an important technological foundation for precision livestock farming. IoT-enabled monitoring systems use environmental sensors, wearable devices, smart collars, cameras, RFID systems, accelerometers, and other connected devices to collect information continuously from livestock and their surrounding environment. These technologies reduce dependence on manual observation and enable remote monitoring of large animal populations.

Environmental sensing units can continuously measure temperature, humidity, harmful gases, air quality, and other housing parameters. Wearable sensors can simultaneously capture animal-specific variables such as body temperature, activity, movement, feeding patterns, and rumination [13]. Communication technologies transfer these measurements to edge or cloud platforms for storage and analysis.

The major advantage of IoT-based monitoring is the availability of high-frequency, multidimensional data. However, raw sensor measurements alone provide limited decision-making value unless they are transformed into meaningful information. Sensor errors, missing observations, communication failures, noise, heterogeneous sampling rates, and large data volumes also create preprocessing challenges. Intelligent analytical techniques are therefore necessary to convert continuously collected livestock data into predictions, risk indicators, and actionable recommendations.

2.4 Machine Learning for Environmental Stress Prediction

Machine learning provides a data-driven mechanism for identifying complex relationships between environmental variables and livestock stress responses. Traditional statistical approaches generally depend on predefined thresholds or linear relationships [14]. In contrast, ML models can learn nonlinear interactions among temperature, humidity, air quality, animal characteristics, behavioral responses, and historical stress conditions.

Algorithms such as logistic regression, decision trees, random forests, support vector machines, gradient boosting, and artificial neural networks have been investigated for classification and prediction tasks in livestock management. Environmental measurements can be combined with animal-specific features to classify conditions into categories such as normal, moderate stress, and high stress. Regression models may alternatively estimate continuous stress indicators or physiological responses.

Feature engineering plays an important role in such models. Derived variables such as THI, moving averages, environmental exposure duration, daily temperature variations, activity changes, and deviations from individual baseline behavior may provide stronger predictive information than isolated sensor readings. Ensemble learning methods are particularly useful when relationships among environmental and physiological parameters are nonlinear [15].

Despite their potential, many environmental monitoring studies remain dependent on predetermined threshold values. Threshold-based systems are simple and interpretable but may not adequately capture variations between animals, breeds, geographical regions, housing systems, or seasons. Adaptive ML-based prediction offers an opportunity to address these limitations by learning risk patterns directly from historical observations.

2.5 AI-Based Disease Risk Prediction

Artificial intelligence has increasingly been applied to livestock disease detection and health-risk prediction. ML models can combine historical health records, physiological measurements, behavioral indicators, production information, and environmental variables to recognize patterns associated with emerging health abnormalities [16]. Classification models can estimate whether an animal belongs to healthy or at-risk categories, while probabilistic approaches can generate disease-risk scores.

Tree-based ensemble models such as random forests and gradient boosting are useful for livestock health applications because they can process heterogeneous features and identify nonlinear interactions [17]. Support vector machines can perform effectively in high-dimensional classification problems, while deep neural networks can learn complex representations from large datasets. Recurrent architectures and other temporal learning methods can further analyze sequential sensor data, allowing changes in health indicators to be evaluated over time.

Nevertheless, disease-prediction models often emphasize physiological or behavioral data without adequately representing environmental exposure. Similarly, environmental stress systems may identify unfavorable climatic conditions without estimating their potential implications for animal health. An integrated model linking environmental stress with disease susceptibility could therefore provide more comprehensive risk assessment.

2.6 Research Gaps and Need for an Integrated Predictive Framework

The literature demonstrates considerable progress in environmental sensing, IoT-based livestock monitoring, machine learning, and automated health assessment. However, several limitations remain. Many existing approaches focus either on environmental monitoring or disease detection rather than establishing an integrated relationship between environmental stress and disease risk. Furthermore, threshold-based monitoring may not capture complex interactions among environmental, behavioral, physiological, and animal-specific variables.

Another limitation is that many monitoring systems primarily provide descriptive information about current conditions rather than predictive information about emerging risks. Although continuous sensor monitoring improves situational awareness, sustainable livestock management requires sufficient lead time for preventive intervention. Prediction of future or emerging stress conditions can allow farmers to modify ventilation, cooling, feeding, water availability, housing conditions, or veterinary monitoring before significant deterioration occurs.

There is therefore a need for an integrated intelligent framework that combines environmental conditions, animal characteristics, behavioral indicators, physiological measurements, and health information within a unified predictive pipeline. Such a framework should perform data preprocessing and feature engineering, estimate environmental stress, evaluate associated disease risk, and convert predictions into interpretable risk levels and management recommendations. Addressing these gaps forms the basis of the proposed intelligent prediction framework for sustainable livestock systems.

3. Proposed Intelligent Prediction Framework

The proposed framework is designed to provide an integrated mechanism for predicting environmental stress and disease risk in livestock systems. It combines environmental conditions with animal-specific, behavioral, and physiological information and transforms these heterogeneous observations into actionable risk estimates using machine learning. Unlike conventional monitoring approaches that primarily identify abnormal conditions after predefined thresholds are exceeded, the proposed framework emphasizes early prediction and preventive decision support.

3.1 Overall Framework Architecture

The proposed framework follows a data-driven architecture in which livestock and environmental information is collected continuously and processed through interconnected analytical stages. Environmental sensors provide information about the surrounding livestock environment, while wearable devices and farm records provide animal-level behavioral, physiological, and health information. These heterogeneous observations are integrated into a common dataset after preprocessing and synchronization.

The overall processing sequence can be represented as:

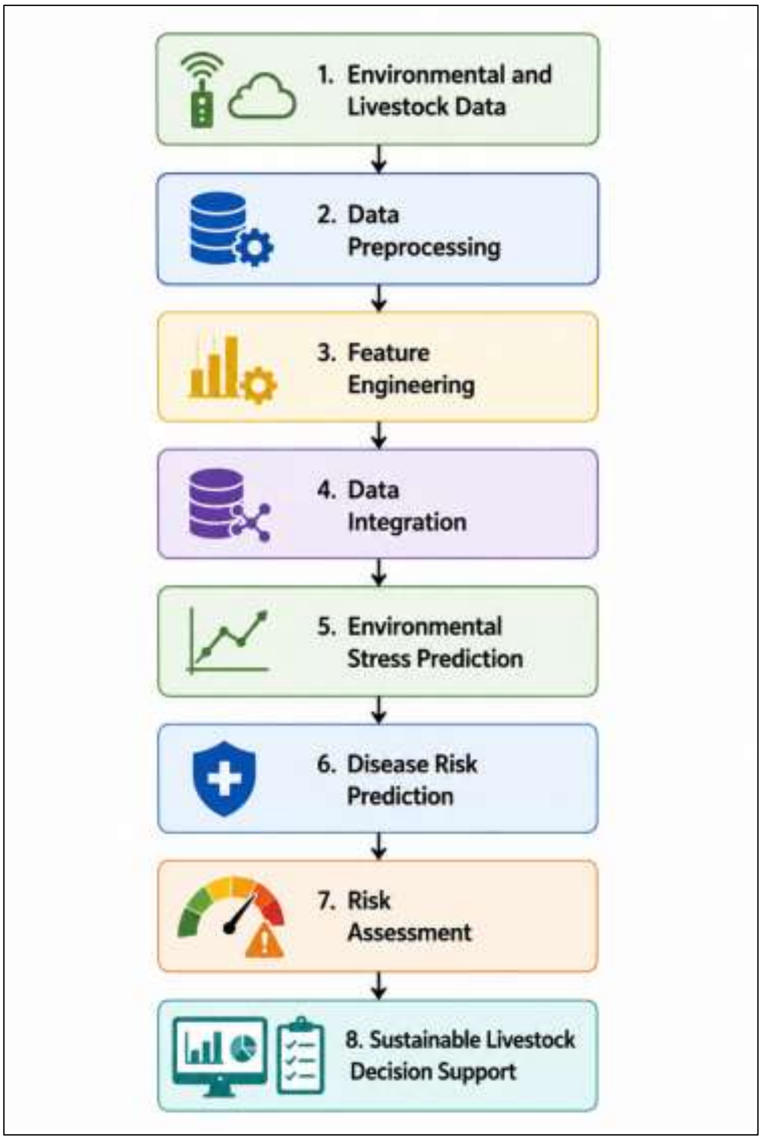

Figure 1. Overall framework for proposed model

The environmental stress prediction stage estimates whether existing conditions are likely to produce physiological stress. The disease-risk prediction stage subsequently considers environmental exposure together with animal-specific indicators to estimate potential health risks. Combining these outputs enables the framework to distinguish between normal environmental variations and conditions requiring management intervention.

3.2 Livestock and Environmental Data Acquisition

Data acquisition represents the first stage of the proposed framework. Multiple sources are considered to provide a comprehensive representation of the relationship between the livestock environment and animal health. Environmental parameters may include temperature, relative humidity, air quality, ammonia concentration, carbon dioxide concentration, ventilation conditions, and other microclimatic variables available within the livestock facility.

Animal-level information can include body temperature, heart rate, respiration rate, activity level, feeding behavior, water consumption, rumination, and movement patterns. Historical health records, disease occurrence, age, breed, weight, and production-related information can also be incorporated when available. The combination of these variables allows the prediction model to evaluate both external environmental exposure and internal animal responses.

The collected data can be represented for an animal at time (t) as:

$$X_t = E_t, B_t, P_t, H_t$$

where (E_t) represents environmental variables, (B_t) represents behavioral indicators, (P_t) denotes physiological measurements, and (H_t) represents historical health and animal-specific information. This multidimensional representation provides the input for subsequent preprocessing and predictive analysis.

3.3 Data Preprocessing and Feature Engineering

Sensor-based livestock datasets may contain missing observations, noisy measurements, duplicated records, inconsistent sampling intervals, and variables measured at different numerical scales. Appropriate preprocessing is therefore necessary before model development. Missing values are handled using suitable interpolation or statistical imputation techniques depending on the nature of the variable. Outliers generated by sensor malfunction are identified and treated to reduce their influence on prediction.

Continuous variables are normalized or standardized when required by the selected machine learning algorithms. Timestamped environmental and animal measurements are synchronized so that physiological and behavioral responses can be associated with corresponding environmental exposure.

Feature engineering is subsequently applied to extract informative patterns from the processed data. In addition to directly measured parameters, derived environmental indicators can be generated. A commonly used measure of thermal stress is the Temperature-Humidity Index (THI), which combines temperature and relative humidity into a single indicator. A generalized form can be expressed as:

$$\mathrm{THI} = (1.8\mathrm{T} + 32) - (0.55 - 0.0055\mathrm{RH})(1.8\mathrm{T} - 26)$$

where (T) represents ambient temperature in degrees Celsius and (RH) represents relative humidity in percentage.

Temporal features such as hourly averages, daily maxima and minima, moving averages, exposure duration, and rate of environmental change can also be generated. Similarly, deviations from normal feeding, activity, body temperature, or rumination patterns can provide early evidence of animal stress. The resulting feature vector therefore contains both raw observations and derived indicators capable of improving predictive performance.

3.4 Environmental Stress Prediction Model

The environmental stress prediction module determines the likelihood that an animal is experiencing or approaching an unfavorable stress condition. Instead of relying exclusively on a fixed environmental threshold, the proposed approach considers multiple environmental and animal-response variables simultaneously.

Let the processed feature vector for environmental stress assessment be represented by $(\mathrm{X_s})$. The stress prediction model can be expressed as:

$$\mathrm{S} = \mathrm{f_s}(\mathrm{X_s})$$

where (f_s) represents the trained machine learning model and (S) represents the predicted environmental stress level. Depending on the application, (S) may be represented as a probability or categorized into classes such as Normal, M8oderate Stress, and High Stress. Multiple machine learning algorithms can be evaluated, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, Gradient Boosting, and Artificial Neural Network models. Comparing different algorithms allows identification of the model that provides the most reliable prediction for the available livestock dataset.

The predicted stress probability can be represented as:

$$\mathrm{P}(\mathrm{S} = 1|\mathrm{X_s})$$

where (S=1) represents the presence of a significant environmental stress condition. Probability-based outputs are particularly useful because they allow risk thresholds to be adjusted according to livestock species, farm conditions, and management requirements.

3.5 Disease Risk Prediction Model

Environmental stress does not necessarily indicate the presence of disease; however, prolonged or severe stress may increase vulnerability to health problems. The second prediction module therefore evaluates disease risk by combining environmental stress information with behavioral, physiological, and historical health indicators.

The disease-risk model can be represented as:

D=f_d(E,B,P,H,S)

where (D) represents predicted disease risk, (E) represents environmental information, (B) denotes behavioral features, (P) represents physiological indicators, (H) denotes historical health information, and (S) represents the estimated environmental stress level.

The resulting model can classify livestock into risk categories such as Low Risk, Moderate Risk, and High Risk. Alternatively, the model can generate a probability:

$$\mathrm{P}(\mathrm{D} = 1|\mathrm{E}, \mathrm{B}, \mathrm{P}, \mathrm{H}, \mathrm{S})$$

where (D=1) represents the occurrence or elevated likelihood of a target health condition.

Including the predicted stress level as an additional input enables the framework to explicitly establish the relationship between environmental exposure and potential health deterioration. For example, persistent high thermal stress accompanied by reduced activity, abnormal body temperature, and reduced feed intake may produce a substantially higher disease-risk estimate than an isolated increase in environmental temperature.

3.6 Sustainability-Oriented Decision Support

The final stage transforms predictive outputs into practical livestock-management information. Environmental stress probability and disease-risk probability are jointly evaluated to determine the urgency and type of intervention required. Rather than presenting farmers with large volumes of raw sensor measurements, the framework generates interpretable risk levels, alerts, and recommended actions. Under normal conditions, routine monitoring can continue without intervention. Moderate environmental stress may trigger recommendations such as improving ventilation, increasing water availability, modifying feeding schedules, or activating cooling mechanisms. Persistent high stress combined with abnormal animal indicators may generate a high-risk alert and recommend closer animal observation or veterinary assessment. The framework therefore supports a transition from reactive livestock management to predictive and preventive management. Early identification of environmental stress can reduce production losses and unnecessary resource consumption, while early disease-risk estimation can support timely intervention and potentially reduce disease severity. The integration of environmental management, animal welfare, health-risk prediction, and intelligent decision support consequently provides a foundation for more efficient and sustainable livestock production systems.

4. Results and Discussion

4.1 Environmental Stress Prediction Results

The performance of different machine learning models was evaluated for predicting environmental stress conditions using integrated environmental and animal-response features. The comparative results are presented in Table 1.

Table 1. Performance comparison for environmental stress prediction

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	85.42	84.16	82.73	83.44	0.87
Decision Tree	87.18	86.35	85.21	85.78	0.88
Random Forest	93.64	92.87	93.15	93.01	0.95
Support Vector Machine	89.76	89.14	88.52	88.83	0.91
Gradient Boosting	**95.21**	**94.68**	**94.91**	**94.79**	**0.97**
Artificial Neural Network	92.83	92.15	91.76	91.95	0.94

Table 2 shows that all evaluated models provide reasonable performance for environmental stress prediction, with accuracy ranging from 85.42% to 95.21%. Logistic Regression records the lowest accuracy of 85.42%, suggesting that a linear decision boundary is comparatively less effective for representing the complex relationships among temperature, humidity, air quality, THI, exposure duration, and animal-response indicators. Decision Tree and SVM improve the prediction accuracy to 87.18% and 89.76%, respectively. The ensemble models demonstrate stronger predictive capability. Random Forest achieves 93.64% accuracy with an F1-score of 93.01%, indicating effective identification of both normal and stressful environmental conditions. Gradient Boosting provides the best overall performance, achieving 95.21% accuracy, 94.68% precision, 94.91% recall, 94.79% F1-score, and an AUC-ROC of 0.97. The high recall indicates that the model successfully identifies most stress-related observations, which is important for preventing potentially harmful environmental exposure.

4.2 Disease Risk Prediction Results

Disease-risk prediction was evaluated by combining environmental conditions, predicted environmental stress, behavioral indicators, physiological measurements, and available animal health information. Table 2 presents the comparative performance of the evaluated machine learning models.

Table 2. Performance comparison for livestock disease-risk prediction

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	83.75	82.61	81.84	82.22	0.85
Decision Tree	86.34	85.72	84.93	85.32	0.87
Random Forest	92.48	91.76	92.13	91.94	0.94
Support Vector Machine	88.91	88.24	87.65	87.94	0.90
Gradient Boosting	**94.36**	**93.72**	**94.08**	**93.90**	**0.96**
Artificial Neural Network	91.67	90.94	91.38	91.16	0.93

As observed in Table 3, the models achieve disease-risk prediction accuracies between 83.75% and 94.36%. Logistic Regression achieves an accuracy of 83.75%, while Decision Tree and SVM improve the accuracy to 86.34% and 88.91%, respectively. The comparatively moderate performance of these models indicates the complexity of relationships among environmental exposure, behavioral abnormalities, physiological variations, and disease risk. Random Forest demonstrates substantially improved performance, achieving 92.48% accuracy and an F1-score of 91.94%. The ANN also provides competitive performance with 91.67% accuracy. However, Gradient Boosting produces the strongest overall results, with 94.36% accuracy, 93.72% precision, 94.08% recall, 93.90% F1-score, and an AUC-ROC of 0.96. Its high recall is particularly important because failure to identify an animal with elevated disease risk may delay preventive or veterinary intervention.

5. Conclusion and Future Research

This study presented an intelligent machine learning framework for predicting environmental stress and disease risk in sustainable livestock systems. The framework integrates environmental, behavioral, physiological, and health-related information to support early identification of unfavorable conditions. Among the evaluated models, Gradient Boosting achieved the best performance, with 95.21% accuracy for environmental stress prediction and 94.36% accuracy for disease-risk prediction, demonstrating the effectiveness of nonlinear ensemble learning for livestock risk assessment. The proposed framework supports preventive livestock management by translating sensor and animal data into interpretable stress and disease-risk levels. Such predictions can assist farmers in making timely decisions regarding ventilation, cooling, feeding, water management, animal observation, and veterinary intervention, thereby improving animal welfare and resource efficiency.

Future research can validate the framework using larger multi-farm and multi-species datasets under different climatic conditions. Temporal deep learning, explainable AI, IoT-based real-time monitoring, and edge computing can also be incorporated to improve prediction, interpretation, and automated decision support. Overall, intelligent risk prediction provides a promising foundation for more resilient, welfare-oriented, and sustainable livestock production.

References

[1] Niloofar, P.; Francis, D.P.; Lazarova-Molnar, S.; Vulpe, A.; Vochin, M.-C.; Suciu, G.; Balanescu, M.; Anestis, V.; Bartzanas, T. Data-Driven Decision Support in Livestock Farming for Improved Animal Health, Welfare and Greenhouse Gas Emissions: Overview and Challenges. Comput. Electron. Agric. **2021**, 190, 106406.

[2] Thumba, D.A.; Lazarova-Molnar, S.; Niloofar, P. Data-Driven Decision Support Tools for Reducing GHG Emissions from Livestock Production Systems: Overview and Challenges. In Proceedings of the 2020 7th International Conference on Internet of Things: Systems, Management and Security (IOTSMS), Paris, France, 14 December 2020; pp. 1–8.

[3] Maze, M.; Taqi, M.O.; Tolba, R.; Abdel-Wareth, A.A.A.; Lohakare, J. Estimation of Methane Greenhouse Gas Emissions from Livestock in Egypt during 1989 to 2021. Sci. Rep. **2024**, 14, 14992.

[4] Chen, C.; Zhu, W.; Oczak, M.; Maschat, K.; Baumgartner, J.; Larsen, M.L.V.; Norton, T. A Computer Vision Approach for Recognition of the Engagement of Pigs with Different Enrichment Objects. Comput. Electron. Agric. **2020**, 175, 105580.

[5] S. N. Ajani, Z. F. Khan, and M. Atique, "AI-driven symptom analysis in veterinary healthcare: A systematic review of machine learning techniques for animal species, breed, and disease identification," *Journal of Animal Environment*, vol. 18, no. 1S, 2026.

[6] Cabrera, V.E.; Niu, M. From isolated data to integrated ecosystems: The artificial intelligence revolution in precision livestock farming. Anim. Front. **2026**, 16, 3–6.

[7] Fuentes, S.; Gonzalez Viejo, C.; Tongson, E.; Dunshea, F.R. The livestock farming digital transformation: Implementation of new and emerging technologies using artificial intelligence. Anim. Health Res. Rev. **2022**, 23, 59–71.

[8] Gorczyca, M.T.; Gebremedhin, K.G. Ranking of Environmental Heat Stressors for Dairy Cows Using Machine Learning Algorithms. Comput. Electron. Agric. **2020**, 168, 105124.

[9] Gorczyca, M. Machine Learning Applications for Monitoring Heat Stress in Livestock. Ph.D. Thesis, Cornell University, Ithaca, NY, USA, 2020.

[10] Rebez, E.; Sejian, V.; Silpa, M.; Kalaignazhal, G.; Thirunavukkarasu, D.; Devaraj, C.; Nikhil, K.; Ninan, J.; Sahoo, A.; Lacetera, N.; et al. Applications of Artificial Intelligence for Heat Stress Management in Ruminant Livestock. Sensors **2024**, 24, 5890.

[11] Koltes, J.E.; Koltes, D.A.; Mote, B.E.; Tucker, J.; Hubbell, D.S., 3rd. Automated collection of heat stress data in livestock: New technologies and opportunities. Transl. Anim. Sci. **2018**, 2, 319–323.

[12] Nowak, P.; Costa, H.; Horvath, I. Smart livestock monitoring using IoT and biosensor technologies. Indo-Am. J. Agric. Vet. Sci. **2025**, 13, 1–9.

[13] Sanjuan, S.; Méndez, D.A.; Arnau, R.; Calabuig, J.M.; Díaz De Otálora Aguirre, X.; Estellés, F. Soft Computing Approaches for Predicting Shade-Seeking Behavior in Dairy Cattle Under Heat Stress: A Comparative Study of Random Forests and Neural Networks. Mathematics **2025**, 13, 2662.
[14] Ghassemi Nejad, J.; Ju, M.-S.; Jo, J.-H.; Oh, K.-H.; Lee, Y.-S.; Lee, S.-D.; Kim, E.-J.; Roh, S.; Lee, H.-G. Advances in methane emission estimation in livestock: A review of data collection methods, model development and the role of AI technologies. Animals **2024**, 14, 435.
[15] Inam, S.A. A review of artificial intelligence for predicting climate driven infectious disease outbreaks to enhance global health resilience. Discov. Public Health **2025**, 22, 738.
[16] Mkhwebu, L.; Dube, S. Artificial Intelligence Techniques in Enhancing Livestock Adaptation to Climate Change: A Systematic Literature Review. Eng. Technol. J. **2026**, 11, 9454–9461.
[17] Quillon Yaprakli, A Comprehensive Review of Environmental Weather Monitoring and Prediction System Using IoT and Multi-Model Progressive Dense Self-Attention. International Journal of Recent Advances in Engineering and Technology, 2023, 12(1), 67 – 72.