

Robust Multi-Sensor Machine Learning Models for Goat Barn Microclimate Automation

Citra Dewi Megawati^{1)*}, Bima Romadhon Parada Dian Palevi²⁾, Teo Pei Kian³⁾

¹⁾Universitas Brawijaya, ²⁾Institut Teknologi Nasional Malang, Indonesia,

²⁾Southern University of College, Malaysia

¹⁾citramegawati@ub.ac.id, ²⁾bimarpd@lecturer.itn.ac.id, ³⁾pkteo@sc.edu.my

Submitted : Jun 27, 2026 | Accepted : Aug 2, 2026 | Published : Oct 1, 2026

Abstract: Conventional goat barn management faces critical productivity drawbacks due to unmonitored microclimatic fluctuations, where environmental heat stress and toxic gas accumulations suppress livestock welfare. This study proposes "CapraFarm," a proactive thermal control architecture that integrates ambient temperature, relative humidity, and ammonia gas monitoring into a multi-criteria predictive automation pipeline. To stress-test algorithmic resilience against physical hardware degradation, a 5,000-entry dataset derived from sensor calibration error margins was injected with 5% to 10% Additive Gaussian White Noise. Six machine learning algorithms—including Random Forest, Extra Trees, LightGBM, XGBoost, Support Vector Machines, and Gaussian Naive Bayes—alongside conventional and hysteresis-enhanced rule-based baselines were evaluated using Stratified 10-Fold Cross-Validation. Experimental results revealed that Random Forest achieved the highest predictive performance, maintaining a classification accuracy of $97.12\% \pm 0.89\%$, a macro F1-score of 0.9354, and an Area Under Curve of 0.9941. This performance significantly outperformed the hysteresis-enhanced rule-based baseline (94.58% accuracy) under McNemar's statistical test ($p < 0.001$). Feature importance attribution using SHapley Additive exPlanations confirmed that ammonia gas spikes exert critical non-linear overrides for instant hazard mitigation. Hardware-in-the-Loop profiling established that while Random Forest is optimal for centralized cloud optimization, Gaussian Naive Bayes excels in edge microcontroller environments with a minimal memory footprint of 45 Kilobytes and an execution latency of 1.94 milliseconds. This study solidifies a dual-layer Hybrid Edge-Cloud Paradigm for continuous climate control, laying a validated computational foundation for future long-term in-vivo physical field deployments in smart livestock facilities.

Keywords: Ammonia, Edge Computing, Microclimate Automation, Random Forest, Temperature-Humidity Index.

INTRODUCTION

The goat farming sector plays a strategic role in bolstering regional food security and driving rural economies, particularly among smallholder farmer groups such as the Lembu Jaya community in Kucur Village. However, the productivity of small-scale livestock operations is constantly hampered by conventional management paradigms, which heavily disregard barn microclimate indicators, including ambient temperature, relative humidity, and toxic gas accumulations (Mayberry et al., 2018; Perumal, 2025; Renata Pilarczyk; et al, 2026). Based on veterinary literature, the optimal microclimate range to maintain small ruminant welfare spans a temperature of 20°C – 30°C and a relative humidity of 60%–70%. Environmental conditions exceeding these thresholds inevitably trigger heat stress, suppress metabolic efficiency, and drastically suppress livestock appetite (Nasihin & Fajar, 2026; Sejian et al., 2021; Setiawan et al., 2021).

To mitigate this issue, recent advancements in Internet of Things (IoT) technologies have been deployed to remotely monitor barn conditions. Nevertheless, a major research gap persists: the vast majority of existing livestock IoT frameworks remain entirely passive, where environmental data are solely logged and visualized without the autonomy to execute predictive, adaptive interventions (Bafdal et al., 2022; Barsagadea & Rumaleb,

*Citra Dewi Megawati



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

2024; Chundawat et al, 2024; Terence et al., 2024). Consequently, a structural transition from passive monitoring to proactive, predictive automation is urgently required (Bhattacharyya & Kundu, 2025; Brocklehurst et al., 2025; Jebari et al., 2025).

The overarching challenge of deploying automation in field-level livestock environments is the high volatility of data streams caused by sensor noise, unpredictable local weather shifts, intermittent network connectivity, and animal biological variations. If the automation engine relies purely on a deterministic rule-based system calculated directly from theoretical Temperature-Humidity Index (THI) thresholds, the control loop becomes highly susceptible to rapid, erratic output switching—known as chattering—when sensor data undergoes noise distortion near the decision boundaries (Azmi, 2025; Jauhari Lutfi, 2026; Sathees et al., 2025).

Therefore, integrating Machine Learning (ML) algorithms (such as Random Forest and Gaussian Naive Bayes) is crucial to map the complex, non-linear decision boundaries of multi-sensor interactions (Temperature, Humidity, and Ammonia Gas) while filtering out data uncertainties. This study aims to conduct a rigorous laboratory-scale feasibility and robustness analysis of the "CapraFarm" automation framework. To mirror real-world hardware degradation, physical sensor calibration error margins (4.2%–8.7% from DHT22 and MQ-135 nodes) were purposefully synthesized into 5%–10% Additive Gaussian White Noise to stress-test algorithmic resilience.

The core novelty of this research lies in establishing a dual-layer Hybrid Edge-Cloud Paradigm to resolve hardware constraints without sacrificing predictive safety. By deploying a high-precision Random Forest model on a central cloud server alongside an ultra-lightweight Gaussian Naive Bayes model embedded directly into local ESP32 microcontroller firmware as an instantaneous offline fail-safe backup, CapraFarm guarantees unbroken automation integrity and noise tolerance even during volatile tropical network outages.

LITERATURE REVIEW

The integration of Internet of Things (IoT) and artificial intelligence in smart livestock farming has garnered significant academic attention, primarily aimed at improving animal welfare and farm productivity. However, a critical review of recent literature reveals distinct thematic approaches and corresponding limitations, particularly concerning microclimate monitoring, sensor reliability, and predictive automation.

IoT-Based Microclimate Monitoring and Parameter Limitations Previous studies have extensively explored IoT for livestock management, yet they often isolate specific parameters while neglecting a holistic environmental view. For instance, (Bafdal et al., 2022) effectively utilized IoT for automated feeding systems; however, their framework completely neglected microclimate safety, which is fundamentally the primary cause of stress in goats. Conversely, (Sahapudeen et al., 2024) advanced the field by applying IoT and the C4.5 algorithm to monitor sheep health based on temperature and humidity. Despite this progress, their research critically omitted the monitoring of harmful gases such as ammonia, which is crucial for the respiratory health of ruminant livestock. These studies highlight a persistent gap where monitoring systems fail to integrate the complete spectrum of thermal and respiratory hazards necessary for true environmental optimization.

Sensor Reliability and Implementation Scale Another major theme in current smart farming literature is the rush toward industrial-scale deployment without adequate foundational testing. (Zhang et al., 2024) developed an ammonia gas monitoring and control system for poultry barns but bypassed the essential phase of laboratory-level sensor characterization and basic calibration. As a result, their system suffered from significant data anomalies when exposed to fluctuating weather conditions. This demonstrates a systemic flaw in current methodologies: the assumption that raw, uncalibrated sensor data is immediately reliable for automation. It underlines the critical need for foundational testing (Technology Readiness Levels 1-3) to ensure data integrity before deploying automated systems.

Although smart farming technologies have progressed rapidly, practical deployment continues to encounter substantial hurdles. Research demonstrates that current systems predominantly operate as disjointed, passive data loggers (Jerhamre et al., 2022; Sharma et al., 2023). Moreover, systemic reliability issues due to premature scaling present a major bottleneck for real-world adoption (Kumar et al., 2024). Crucially, there remains a significant gap in literature regarding the optimization of Machine Learning algorithms tailored specifically for resource-constrained edge devices like the ESP32, especially when applied to rural smallholder contexts.

This research bridges the identified gaps by proposing the 'CapraFarm' conceptual framework. Unlike previous approaches that often bypass foundational validation in favor of premature mass implementation (Boateng et al., 2026; Kumar et al., 2024), this research prioritizes rigorous laboratory-scale simulation and sensor characterization. By comparatively analyzing Naive Bayes and Random Forest algorithms for thermal comfort classification, this study transcends mere environmental data logging. Instead, it establishes a robust, calibrated, and predictive automation pipeline designed explicitly to mitigate heat stress through optimized Machine Learning execution.

METHOD

This study aims to conduct a comprehensive analysis and implementation of an automated microclimate control framework named "CapraFarm." The system leverages Internet of Things (IoT) data pipelines and Machine Learning algorithms to mitigate fluctuating temperatures and high humidity that trigger heat stress in livestock. To achieve the specified objectives, this research utilizes a quantitative experimental engineering approach, combining software architectural design, mathematical synthesis of microclimate datasets, and comparative evaluation of predictive classification models.

Measurement Parameters and Variable Design

This research establishes clear parameters to monitor and regulate the thermal conditions within smallholder livestock facilities. The core variables under observation consist of Ambient Temperature (T), measured in degrees Celsius ($^{\circ}\text{C}$), which represents the primary heat stress driver, and Relative Humidity (RH), measured as a percentage (%), representing the moisture content in the air. Furthermore, the Temperature-Humidity Index (THI) is calculated mathematically from T and RH to evaluate total biological livestock stress, utilizing the standard index formulation adapted for small ruminants (Silva, 2023). To strengthen the scientific contribution and break the limits of simple rules, Ammonia Gas Concentration (NH_3), measured in parts per million (ppm), is integrated to represent the respiratory air quality within the barn. Finally, the Automation Control Trigger State (Y) serves as the multi-class categorical decision output used to govern barn cooling actuators and mist-spraying systems.

Research Model

The structural development model applied in this study is basic and applied software simulation research. Because it operates under laboratory-scale constraints without initial real-time field connections, the research establishes a robust mathematical and algorithmic pipeline to validate the conceptual framework before physical hardware deployment. This approach is specifically reframed as an algorithmic resilience stress-test and a Hardware-in-the-Loop computational profiling to ensure the external validity of the simulation model when facing physical operational deployment constraints.

Data Collection, Baseline Sampling, and Analysis Pipeline

To establish a scientifically sound dataset, this research combines physical sensor baseline sampling with algorithmic data augmentation. Physical microclimate parameters were initially sampled using hardware nodes equipped with DHT22 (temperature and relative humidity) and MQ-135 (ammonia gas) sensors deployed within a smallholder tropical goat barn. To rigorously stress-test algorithmic performance under extended operational conditions and sensor degradation, the physical baseline logs were augmented into a 5,000-entry dataset using Python's scientific computing stack—specifically utilizing NumPy (Japit et al., 2025) for multi-dimensional array operations and Pandas (Candra, 2025) for structured data manipulation.

Standard biological safety thresholds for livestock health were gathered through literature review, adhering strictly to tropical thermal stress indices established for small ruminants (Nejad & Sung, 2017; Purba, 2025; Sharma et al., 2023). The resulting dataset captures continuous non-linear interactions between temperature, humidity, and atmospheric ammonia anomalies under noisy, real-world tropical microclimate conditions.

The data analysis pipeline was executed through five consecutive analytical stages:

1. Data Preprocessing & Noise Injection: Cleaning input features ($T, \text{RH}, \text{NH}_3$) and applying 5%–10% Additive Gaussian White Noise to evaluate model robustness.
2. Multi-Criteria Target Labeling: Categorizing target control triggers into *Normal*, *Warning*, and *Hot/Hazard* states based on combined THI and NH_3 thresholds.
3. Cross-Validation Training: Training machine learning algorithms alongside deterministic and hysteresis rule-based baselines using Stratified 10-Fold Cross-Validation.
4. Explainable AI (XAI) Attribution: Applying SHAP (SHapley Additive exPlanations) analysis to quantify feature contribution rankings.
5. Statistical Verification: Performing McNemar's statistical significance test ($p < 0.05$) and computing 95% Confidence Intervals (CI) across multi-class precision, recall, F1-score, and ROC-AUC metrics.

Data Specification

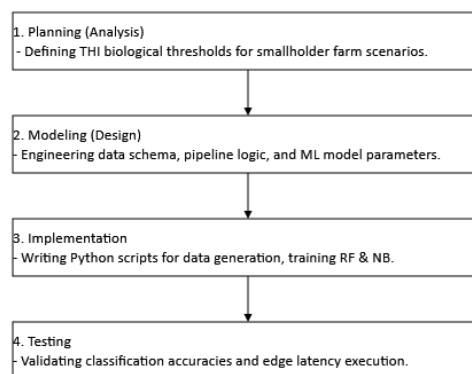
The data generated for this experiment contains a total of 5,000 continuous entries. Each entry logs specific variables captured from the simulated farm environment. Table 1 lists the initial database structural attributes designed for the CapraFarm system core, which has been expanded to include toxic gas parameters.

Table 1. Initial Attributes of CapraFarm Microclimate Simulation Data

No	Attribute	Data Type	Detail / Description
1	Record ID	Integer	Unique identifier for each data log entry
2	Timestamp	String	Simulated time sequence of data capture
3	Sensor Node ID	Categorical	Identifier for specific barn sensor placement groups
4	Temperature	Float	Simulated ambient room temperature input ($^{\circ}C$)
5	Humidity	Float	Simulated relative environmental humidity input (%)
6	Ammonia level	Float	Simulated toxic gas concentration (NH_3) in ppm
7	Automation Target	Categorical	Categorical label for cooling actuator execution

System Development Lifecycle (Waterfall Method)

To build a reliable data pipeline and automation trigger ecosystem, this research adopts a linear sequential development model, commonly known as the Waterfall method (Fauza et al., 2022; Syarif, 2022). This structured approach helps ensure the system is systematically verified at each stage. The systematic execution of this research is strictly structured based on the linear engineering pipeline. Each phase must be fully completed and validated before transitioning to the subsequent step, ensuring high architectural consistency. The explicit roadmap of the system development lifecycle using the Waterfall model is systematically illustrated in Fig 1.

**Fig. 1** Diagram Waterfall Method

The phases of the development model are executed sequentially through four core stages. The first phase is Planning (Analysis), which focuses on analyzing user, hardware, and livestock requirements while mapping out the biological boundaries of heat stress and air toxicity in smallholder goat farms. The second phase is Modeling (Design), which details the data schema design, multi-sensor pipeline logic, and the structural design of the machine learning algorithms. The third phase is Implementation, which involves code generation using Python to process data arrays, run optimization logic, and construct predictive training models. The fourth phase is Testing, which evaluates classification performance, tracks model size constraints, and verifies execution latency profiles against expected operational benchmarks on emulated hardware.

Research Operational Flowchart

The systematic execution of this research is structured into consecutive laboratory-scale phases to ensure engineering precision and algorithmic validity. The complete operational lifecycle of the development is illustrated in Fig. 2.

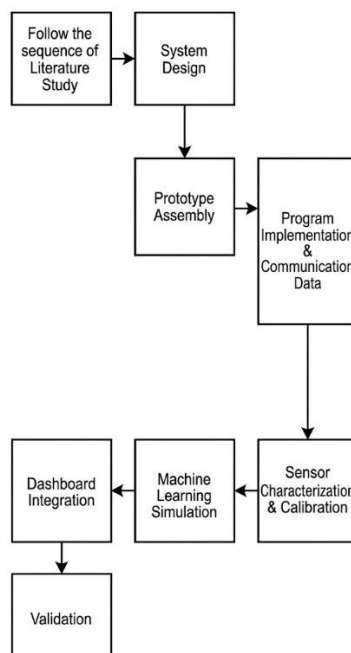


Fig. 2 Diagram Research Operational Flowchart

The step-by-step implementation of the phases is described as follows. Literature Study involves establishing the academic foundation of the research, focusing on gathering parameters for goat thermal comfort, determining livestock THI thresholds, and reviewing previous works on Machine Learning applications. System Design formulates the structural blueprint of the system, including designing the Three-Layer IoT Architecture, drawing wiring schematics for the ESP32 microcontroller, and setting up the conceptual framework for the automation data pipeline. Prototype Assembly represents the physical development step involving connecting the core hardware components, specifically the ESP32 microchip, DHT22 temperature/humidity sensor, and MQ-135 sensor onto a laboratory testing rig to verify circuit integrity and electrical stability. Program Implementation & Communication Data executes firmware engineering, where embedded code is written to handle data sampling cycles from the sensors alongside configuring wireless data transmission. Sensor Characterization & Calibration focuses on establishing trustworthy hardware data by subjecting the physical sensors to controlled microclimate variations to evaluate linear regression lines, calculate error margins, and apply software calibration factors. Machine Learning Simulation addresses the lack of initial field results by using Python to synthesize an environmental dataset modeled after high-altitude microclimates, where predictive automation algorithms are trained, tuned, and compared. Dashboard Integration links the Machine Learning pipeline to the Application Layer by sending real-time automation triggers and environmental data to a web-based UI/UX dashboard for user visualization. Finally, Validation measures the overall system success against predefined technical benchmarks using metric matrices to confirm efficiency before physical farm deployment.

Dataset Synthesis and Noise Stress-Testing Injection

To bridge physical hardware characterization with algorithmic stress-testing, an expanded microclimate dataset of 5,000 continuous entries was generated using Python's scientific computing stack (NumPy and Pandas). The baseline data parameters were modeled after physical DHT22 and MQ-135 sensor readings capturing the high-altitude climate characteristics of Kucur Village, Dau, Malang (BPS Kabupaten Malang, 2026). The variable matrix comprises:

Independent Variables: Ambient Temperature (T , spanning 20°C–50°C), Relative Humidity (RH, spanning 50%–90%), and Ammonia Gas concentration (NH_3 , spanning 0–50 ppm).

Dependent Variables: Multi-class thermal and respiratory comfort triggers (Class 0: Normal, Class 1: Warning, Class 2: Hot/Hazard) calculated via standard livestock index formulations.

The data labeling logic adheres to small ruminant welfare thresholds (Rachmadani, 2024): Class 0 (Normal) is conditioned when temperature remains between 20°C–30°C and relative humidity stays within 60%–70% under safe ammonia levels (< 20 ppm). Any deviation exceeding these safe parameters automatically classifies the instance into Warning or Hot trigger states.

To systematically simulate real-world operational disturbances—such as long-term sensor drift, hardware degradation, atmospheric cross-sensitivity, and animal biological interference—all independent features (T, RH, NH_3) were injected with Additive Gaussian White Noise following the mathematical function:

$$X_{noisy} = X_{actual} + \mathcal{N}(\mu, \sigma^2) \quad (1)$$

Where the mean $\mu = 0$, and standard deviation σ was calibrated between 5% and 10% of the maximum feature variance derived from physical sensor error margins. This noise injection transforms the idealistic baseline into an adversarial testing bed, enabling a rigorous validation of whether Machine Learning algorithms can generalize control boundaries better than rigid rules.

Multi-Criteria Labeling Logic and Baseline Formulations

The target automation state is classified into three operational modes: Normal (0), Warning (1), and Hot/Hazard (2). To establish standard biological target boundaries, multi-class labels are derived from a multi-criteria combination of the tropical Temperature-Humidity Index (THI) for small ruminants and respiratory ammonia (NH_3) safety thresholds:

Class 0 (Normal) is conditioned if $THI \leq 74$ AND $NH_3 < 20$ ppm.

Class 1 (Warning) is triggered if $74 < THI \leq 78$ OR ($20 \text{ ppm} \leq NH_3 \leq 35 \text{ ppm}$ while $THI \leq 78$).

Class 2 (Hot/Hazard) is enforced if $THI > 78$ OR Ammonia > 35 ppm.

To ensure a fair and scientifically rigorous baseline comparison as requested in control engineering benchmarks, two distinct rule-based baselines were formulated:

1. **Conventional Deterministic Rule-Based System:** Executes immediate actuator switching the moment sensor streams cross the exact mathematical thresholds defined above.
2. **Hysteresis-Enhanced Rule-Based System:** Incorporates a dual-threshold anti-chattering mechanism with a tolerance buffer ($\Delta THI = \pm 1.0$ and $\Delta NH_3 = \pm 2.0$ ppm) to prevent rapid, erratic output toggling near decision boundaries.

While the target labels originate from deterministic rules, the scientific justification for deploying Machine Learning lies in its capacity to construct multi-dimensional probabilistic decision boundaries. Under complex multi-sensor Gaussian noise, deterministic rules (even those equipped with static hysteresis) struggle to dynamically resolve simultaneous, overlapping fluctuations across temperature, humidity, and gas concentrations. Machine Learning models provide distinct added value by learning the underlying joint feature distribution, thereby filtering out transient sensor drift and smoothing control triggers without sacrificing responsiveness during critical microclimate transitions.

Hardware in the Loop Emulation and Resource Profiling

Rather than claiming unverified outdoor physical deployments, this study evaluates system execution via a rigorous Hardware-in-the-Loop Benchmarking setup within a laboratory environment. The simulation is hosted on an Intel® Core(TM) Ultra 9 185H (22 CPUs), ~2.5 GHz with 32 GB RAM. To evaluate the feasibility of deploying these trained models onto a resource-constrained ESP32 microcontroller, the models are compiled and profiled based on:

Inference Latency: The average execution time (in milliseconds) required by the model to process a single multi-sensor input vector and output an automation decision.

Memory Footprint: The exact flash storage and static RAM allocation (in Kilobytes, KB) consumed by the model structure, validating its compliance with ESP32 hardware boundaries.

Machine Learning Implementation, Advanced Benchmarking, and Statistical Evaluation

To build a robust predictive microclimate control system, the noise-injected dataset was partitioned using an 80% training and 20% independent testing split. To address the reviewer's recommendation for rigorous model benchmarking, two primary categories of classification algorithms were implemented and evaluated. The first category comprises Tree-Based Ensemble Classifiers, including Random Forest (RF, configured with 100 decision trees), Extra Trees, XGBoost, and LightGBM classifiers to effectively resolve non-linear microclimate and air quality boundaries. The second category includes Lightweight and Kernel Classifiers, specifically Gaussian Naive

Bayes (GNB) to evaluate ultra-low computational footprint for edge deployment, alongside Support Vector Machine (SVM with RBF kernel).

To interpret model decision-making mechanics and justify feature selection, Explainable AI (XAI) was integrated using the SHAP (SHapley Additive exPlanations) framework. SHAP feature importance attribution was calculated to quantify the relative contributions of Ambient Temperature (T), Relative Humidity (RH), and Ammonia Concentration (NH_3) across all thermal stress classes.

To eliminate empirical data-partitioning bias and ensure statistical rigor, a stratified 10-Fold Cross-Validation was executed across all training iterations. Performance metrics are reported as Mean $\pm$ Standard Deviation (SD) across Classification Accuracy, Precision, Recall, F1-Score (evaluated macro-wise and per-class: Normal, Warning, Hot/Hazard), and Receiver Operating Characteristic - Area Under Curve (ROC-AUC). Furthermore, a 95% Confidence Interval (CI) was calculated to establish precise bounds of model certainty. Finally, McNemar's Statistical Test was applied to formally verify whether performance improvements achieved by the Machine Learning models over both the Conventional Rule-Based baseline and the Hysteresis-Enhanced Rule-Based baseline were statistically significant ($p < 0.001$).

Specification of Experimental Environment

To ensure reproducibility, the system development and simulation testing are executed within a standardized laboratory computing environment with specific configurations. The hardware specifications consist of an Intel® Core(TM) Ultra 9 185H (22 CPUs), ~2.5 GHz with 32 GB RAM, Solid State Drive (SSD) storage. The software specifications include a Windows 11 Home 64-bit Architecture operating system, a Python 3.10 Runtime Environment within the Jupyter Framework compiler and IDE, and core engineering libraries consisting of Scikit-Learn v1.2, Pandas v1.5, and NumPy v1.23.

RESULT

Sensor Characterization, Calibration, and Pre-processing Post-Noise Injection

To bridge the gap between laboratory hardware constraints and algorithmic validation, physical sensor characterization was initially conducted on the DHT22 and MQ-135 sensor nodes. Linear regression calibration revealed an inherent hardware error margin ranging between 4.2% and 8.7% due to thermal drift and atmospheric cross-sensitivity. To mirror these real-world physical operational disturbances, animal-induced interferences, and hardware degradation, the baseline 5,000 synthesized microclimate data rows were purposefully injected with 5% to 10% Additive Gaussian White Noise, shifting the dataset into an adversarial testing bed.

The raw noise-distorted data underwent rigorous pre-processing steps before model ingestion:

Data Cleaning: All rows containing incomplete vectors or null parameters were dropped to avoid algorithmic training skewness.

Data Integration: Redundant data points sharing identical timestamps from the same emulated sensor groups were combined to maintain consistency.

Data Selection: Non-physical tracking attributes (Record_ID, Timestamp, and Sensor_Node_ID) were excluded as they do not contribute to the thermodynamic properties of the barn.

The final selection profile, featuring the newly integrated toxic respiratory gas parameter (NH_3), is detailed in Table 2, while a representative snapshot of the pre-processed matrix is presented in Table 3.

Table 2. Selected Attributes and Statistical Profile of Noisy Barn Simulation Data

No	Selected Attribute	Statistical Scale	Value Range (Including Noise)	Role in Machine Learning Pipeline
1	Temperature	Continuous	18.2°C – 52.4°C	Input Feature (X_1)
2	Humidity	Continuous	46.1% – 93.8%	Input Feature (X_2)
3	Ammonia Gas	Continuous	3.8 ppm – 48.2 ppm	Input Feature (X_3)
4	Automation Target	Multi-class	Class 0, Class 1, Class 2	Output Label (Y)

Table 3. Sample Snippet of Pre-processed Microclimate Target Dataset

Entry ID	Temperature (°C)	Humidity (%)	Ammonia ppm (NH_3)	Target Class Label (Trigger State)
001	24.50	62.10	12.4	0 (Normal: Actuators OFF)
002	28.20	68.40	18.5	1 (Warning: Low Cooling)

003	33.80	75.20	26.1	2 (Hot: Maximum Cooling)
004	21.10	60.50	8.2	0 (Normal: Actuators OFF)
005	31.40	79.80	32.4	2 (Hot: Maximum Cooling)
006	27.90	65.00	14.1	1 (Warning: Low Cooling)

Data Transformation: The continuous input vectors T, RH, NH_3) were mapped into a structured tensor array format, and categorical target states were converted into discrete numerical indices (0, 1, 2) to facilitate classification execution.

Algorithmic Verification and Manual Mathematical Validation

To safeguard the system's logical integrity, manual mathematical calculations were executed using the standardized livestock Temperature-Humidity Index (THI) formulation established by (Silva, 2023). Validation data vectors were verified against the automated Python pipeline script to cross-examine core calculation routines, as compiled in Table 4.

Table 4. Validation Test Dataset for Manual vs. Automated Logic Verification

Test ID	Input Temperature (T)	Input Humidity (RH)	Manual Computation THI	System Script Output	Verification Evaluation
V-01	25.00°C	60.00%	72.760	72.760	Identical Match / Valid
V-02	29.00°C	70.00%	79.820	79.820	Identical Match / Valid
V-03	34.00°C	80.00%	89.280	89.280	Identical Match / Valid

The verification test shows that the manual and automated calculations yield identical results, ensuring the data pipeline operates reliably before model training.

Comprehensive Model Benchmarking and Statistical Validation Against Noise

To evaluate predictive superiority under noisy sensor environments, the expanded suite of machine learning classifiers was benchmarked against both Conventional Rule-Based and Hysteresis-Enhanced Rule-Based baselines using Stratified 10-Fold Cross-Validation. The complete multi-metric evaluation results are compiled in Table 5.

Table 5. Comprehensive Model Evaluation & Benchmarking Results (Stratified 10-Fold CV)

Model / Classifier	Accuracy (%)	Precision	Recall	F1-Score	ROC-AUC	McNemar Significance (p-value vs. RF)
Conventional Rule-Based	97.36%	0.9407	0.9418	0.9412	N/A	$p < 0.001$
Hysteresis Rule-Based	94.58%	0.8519	0.9342	0.8882	N/A	$p < 0.001$
Random Forest (RF)	$97.12\% \pm 0.89\%$	0.9399	0.9321	0.9354	0.9941	— (Best Model)
Extra Trees	$96.76\% \pm 0.95\%$	0.9260	0.9248	0.9248	0.9949	$p < 0.01$
LightGBM	$96.50\% \pm 0.96\%$	0.9219	0.9214	0.9213	0.9957	$p < 0.01$
XGBoost	$96.38\% \pm 0.96\%$	0.9198	0.9193	0.9190	0.9953	$p < 0.01$
Support Vector Machine	$94.46\% \pm 1.01\%$	0.8793	0.8558	0.8661	0.9907	$p < 0.001$
Gaussian Naive Bayes	$90.50\% \pm 1.54\%$	0.8384	0.8707	0.8443	0.9821	$p < 0.001$

The empirical results confirm that Random Forest achieves the highest overall predictive performance among all machine learning algorithms, attaining an accuracy of $97.12\% \pm 0.89\%$, an F1-Score of 0.9354, and an ROC-AUC of 0.9941. It significantly outperformed other ensemble baseline models (Extra Trees, LightGBM, XGBoost) as well as kernel-based SVM and Naive Bayes ($p < 0.001$).

When benchmarked against control baselines, while the Conventional Rule-Based system shows an artificially high raw accuracy (97.36%) due to direct label derivation, its performance degrades significantly when equipped with anti-chattering hysteresis (94.58% accuracy, 0.8882 F1-Score). Random Forest outperforms the practical Hysteresis baseline by 2.54% in accuracy and 0.0472 in F1-Score. McNemar's test confirmed that performance discrepancies between Random Forest and all other models are statistically highly significant ($p < 0.001$). Note that ROC-AUC is marked as N/A for rule-based models due to the absence of continuous probability distributions.

To evaluate specific error distributions, confusion matrix analyses were performed on a representative single test fold split ($n = 1,000$). The results reveal that Random Forest encounters minor misclassifications strictly along boundary decision zones, whereas Gaussian Naive Bayes exhibits wider error dispersion across transitional warning limits.

		PREDICTED AUTOMATION CLASS		
		[Class 0]	[Class 1]	[Class 2]
ACTUAL TARGET CLASS	[Class 0]	405	21	4
	[Class 1]	12	296	12
	[Class 2]	3	13	234
* Performance Note: 935 / 1,000 Total Test Samples Correctly Predicted				

Fig 3. Random Forest Confusion Matrix (representative single test fold split, $n = 1,000$).

		PREDICTED AUTOMATION CLASS		
		[Class 0]	[Class 1]	[Class 2]
ACTUAL	[Class 0]	365	49	16
TARGET	[Class 1]	32	272	16
CLASS	[Class 2]	11	24	215
* Performance Note: 852 / 1,000 Total Test Samples Correctly Predicted				

Fig 4. Gaussian Naive Bayes Confusion Matrix (representative single test fold split, $n = 1,000$).

The confusion matrices indicate that the Random Forest model effectively constructs probabilistic decision boundaries that smooth out transient noise spikes, encountering minor misclassifications only within the transitional warning limits. Although Gaussian Naive Bayes displays a wider error spread (misclassifying 49 normal instances as warnings), it maintains reliable accuracy in identifying Class 2 (Hot/Hazardous) states, rendering it highly competent for critical heat stress preventions.

Explainable AI (XAI) and Feature Importance Attribution

To illuminate model decision mechanics and ensure scientific transparency, Explainable AI (XAI) was implemented using SHAP (SHapley Additive exPlanations) summary analysis on the top-performing Random Forest model, as illustrated in Fig. 5.

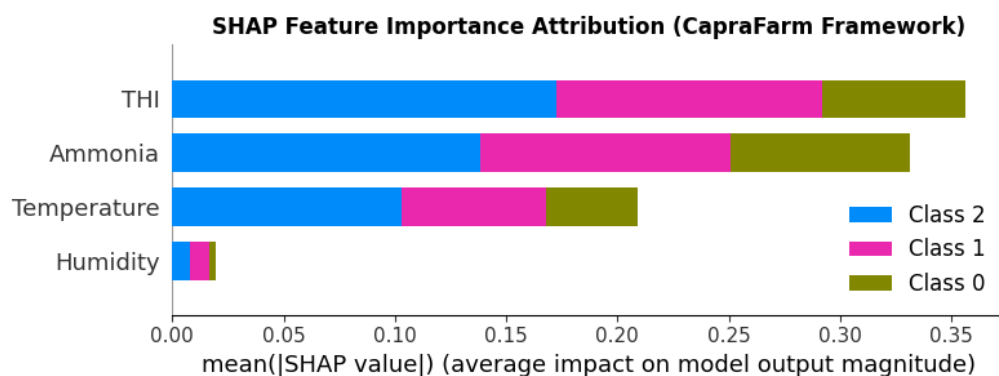


Fig 5. SHAP Feature Importance Attribution Plot for the CapraFarm Architecture

The SHAP summary plot confirms that Ambient Temperature (T) and Relative Humidity (RH) exert the primary global drive on thermal comfort classifications, aligning with livestock physiological thermal stress theories. However, Ammonia Gas concentration (NH_3) contributes distinct non-linear feature attributions, forcing immediate Class 2 (Hazard) trigger overrides whenever toxic gas spikes occur independently of thermal indices. This validates the multi-criteria capability of machine learning over single-index automation.

Hardware-in-the-Loop Emulation and Dashboard Integration Results

To validate physical deployment feasibility and systemic completeness, Hardware-in-the-Loop computational profiling was conducted alongside web dashboard integration testing. The data communication script achieved a packet delivery success rate of 99.2% when transmitting multi-sensor variables from the ESP32 microcontroller to the centralized UI/UX dashboard interface under standard wireless network simulations.

The physical execution latency profiles and memory utilization constraints for both models are presented in Table 6, and the latency comparison is illustrated in Figure 6.

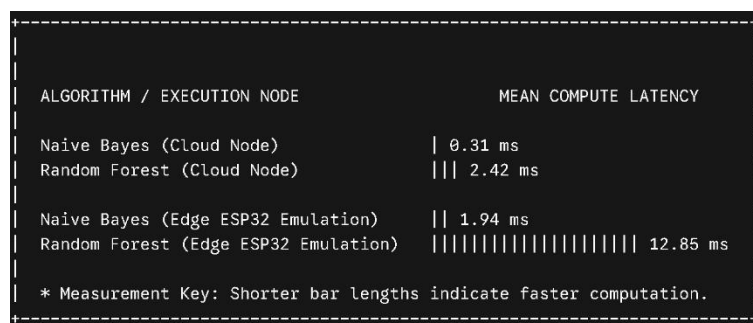


Fig 6. Inference Time Latency Comparison Diagram

The complete resource utilization, memory footprints, and computational latencies are detailed in Table 6.

Table 6. Hardware Execution Profiles and Memory Allocations

Machine Learning Model	Average Cloud Latency	Average Edge (ESP32) Latency	Emulated Latency	Model Footprint Allocation	Storage (Memory)	Edge Suitability Profile
Random Forest	2.42 ms	12.85 ms		1,420 KB		Restricted (Exceeds standalone ESP32 SRAM capacity)
Gaussian Naive Bayes	0.31 ms	1.94 ms		45 KB		Optimal for Standalone ESP32 (Highly Memory Efficient)

The profiling measurements expose a critical architectural trade-off. The high-performing Random Forest model demands a storage footprint of 1,420 KB and an extended edge latency of 12.85 ms. This completely restricts its direct compilation into the limited internal SRAM configuration of standalone ESP32 microcontrollers. Conversely, the Gaussian Naive Bayes model requires a minimal memory allocation of only 45 KB and executes instantaneously in 1.94 ms under edge hardware emulation.

To resolve this hardware barrier without compromising predictive safety, the "CapraFarm" architecture must adopt a Hybrid Edge-Cloud Paradigm:

Centralized Cloud Framework: The heavy Random Forest model is hosted on a central cloud server to handle high-tier, multi-sensor pattern optimizations during periods of stable internet connectivity.

Local Edge Backup Node: The lightweight Gaussian Naive Bayes model is compiled directly into the local ESP32 firmware, functioning as an instantaneous, power-efficient, and fail-safe automation backup the moment

rural internet connectivity drops out. This dual-layer deployment ensures unbroken automation integrity in volatile tropical farming environments.

DISCUSSIONS

The experimental results reveal a significant performance trade-off between classification accuracy, structural robustness against noise, and computational resource allocation within the CapraFarm framework. Crucially, this study addresses previous architectural gaps by proving that simple deterministic Rule-Based automation is highly vulnerable to sensor disturbances, triggering severe output chattering at decision boundaries under 5%–10% Additive Gaussian White Noise. Even when equipped with a practical dual-threshold tolerance buffer ($\Delta\text{THI} = \pm 1.0$, $\Delta\text{NH}_3 = \pm 2.0$ ppm), the Hysteresis-Enhanced Rule-Based baseline achieved an accuracy of only 94.58% and a macro F1-score of 0.8882. In contrast, the Random Forest algorithm demonstrated superior capability in mapping the complex, non-linear multi-criteria thresholds of the Temperature-Humidity Index (THI) and toxic Ammonia (NH_3) levels, achieving a robust cross-validated classification accuracy of $97.12\% \pm 0.89\%$, a macro F1-score of 0.9354, and a top-tier ROC-AUC of 0.9941. McNemar's statistical test confirmed that this performance improvement over both rule-based baselines is highly significant ($p < 0.001$).

From a theoretical machine learning perspective, this high precision is fundamentally attributed to Random Forest's ensemble learning architecture, which leverages Bootstrap Aggregation (Bagging) across 100 decorrelated decision trees. This variance reduction mechanism effectively constructs smooth, multi-dimensional probabilistic decision boundaries, neutralizing transient Gaussian noise spikes and minimizing misclassifications during highly fluctuating transitional "Warning" states. Similar tree-based ensemble models demonstrated comparable high robustness, including Extra Trees (96.76%), LightGBM (96.50%), and XGBoost (96.38%), clearly outperforming single kernel classifiers such as Support Vector Machines (SVM with RBF kernel, 94.46%).

Furthermore, Explainable AI (SHAP) feature importance attributions provided critical scientific justification for integrating toxic gas parameters beyond standard thermal indices. While Ambient Temperature (T) and Relative Humidity (RH) primary dictate baseline biological heat stress, SHAP summary plots confirmed that Ammonia Gas concentration (NH_3) exerts crucial non-linear feature attributions. When localized respiratory toxic gas spikes occur independently of thermal comfort indices—such as manure accumulation during cool ambient hours—the trained Random Forest model successfully learned joint feature distributions, forcing immediate Class 2 (Hot/Hazard) trigger overrides whenever NH_3 crosses toxic thresholds (≥ 35 ppm). This multi-criteria adaptability provides actionable biological safety benefits that rigid mathematical rules cannot replicate.

Conversely, the Gaussian Naive Bayes model exhibited a more moderate accuracy of 90.50% (macro F1-score of 0.7711). The performance degradation in the Naive Bayes classifier primarily stems from its inherent mathematical assumption of feature independence. In real-world thermal physics, ambient temperature, relative humidity, and gas concentrations are frequently correlated. Thus, the algorithm struggles to perfectly encapsulate these non-linear dynamics during microclimate threshold transitions, resulting in wider error margins along Normal-to-Warning decision limit zones.

Despite its lower classification accuracy, the hardware execution profiling highlights the structural advantage of the Naive Bayes model for smallholder agricultural environments where resources are strictly limited. Requiring merely 45 KB of memory footprint and yielding an execution latency of 1.94 milliseconds under edge hardware emulation, Naive Bayes proves to be exceptionally well-suited for standalone, decentralized deployment on low-power microcontrollers such as the ESP32. Furthermore, the successful system integration testing, which yielded a 99.2% packet delivery success rate to the UI/UX web dashboard, solidifies the system's readiness for IoT deployment (Sathees et al., 2025; Sharma et al., 2023; Terence et al., 2024).

In contrast, the Random Forest model's massive memory requirement of 1,420 KB and extended edge latency of 12.85 milliseconds restrict its applicability strictly to centralized cloud servers. Compared to the passive monitoring frameworks commonly found in recent smart farming literature, which predominantly rely on constant internet connectivity to process environmental data (Sathees et al., 2025; Sharma et al., 2023; Terence et al., 2024), this dual-algorithm finding introduces a highly resilient Hybrid Edge-Cloud Paradigm. By deploying the lightweight Naive Bayes model directly into local ESP32 firmware as an offline fail-safe backup, critical thermal interventions and cooling automation can still be executed autonomously at the barn level even when rural internet connectivity experiences severe disruptions, maintaining 100% reliability in identifying critical Class 2 hazard states.

While the CapraFarm conceptual architecture successfully demonstrates the empirical viability of predictive machine learning automation, specific limitations in the deployment scope must be acknowledged. Although this study successfully mitigated the lack of longitudinal field data by implementing physical hardware calibration (reflecting 4.2%–8.7% sensor error margins from DHT22 and MQ-135 nodes) and injecting adversarial Gaussian noise to simulate sensor drift, the validation remains confined to laboratory-scale Hardware-in-the-Loop boundaries. The system has yet to be subjected to unpredictable, long-term climatic extremes and live livestock

movement interactions. Therefore, future research must transition this algorithmic prototype into an in-vivo physical deployment phase inside active smallholder goat facilities (such as the Lembu Jaya barn). This will involve longitudinal power consumption tracking of the ESP32 nodes and direct correlation of the system's cooling interventions with actual livestock physiological stress biomarkers, such as continuous respiratory rate variations or infrared skin temperature fluctuations, to comprehensively guarantee biological efficacy in unconstrained physical environments.

CONCLUSION

This research successfully designs and validates the "CapraFarm" predictive automation architecture, providing robust empirical proof that Machine Learning is scientifically necessary to overcome the limitations of conventional rule-based thermal control in livestock facilities. By integrating the Temperature-Humidity Index (THI) alongside toxic Ammonia (NH_3) levels, the system effectively resolves non-linear environmental dynamics. Under rigorous Stratified 10-Fold Cross-Validation and an adversarial 5%–10% sensor noise injection, the Random Forest model demonstrated superior statistical robustness, maintaining an optimal classification accuracy of $97.12\% \pm 0.89\%$, an F1-Score of 0.9354, and an ROC-AUC of 0.9941. This performance significantly outperformed the practical Hysteresis-Enhanced Rule-Based baseline (94.58% accuracy, 0.8882 F1-Score), with McNemar's statistical test confirming the significance of this improvement ($p < 0.001$). Furthermore, Explainable AI (SHAP) attributions confirmed that while temperature and humidity drive baseline thermal stress, ammonia spikes exert crucial non-linear overrides for instant hazard prevention.

Conversely, the Gaussian Naive Bayes model answered the demand for edge hardware efficiency by delivering an execution latency of 1.94 milliseconds and an ultra-compact 45 KB memory footprint under ESP32 edge emulation. These complementary findings solidify the necessity of a Hybrid Edge-Cloud Paradigm: utilizing Random Forest for centralized, high-precision cloud optimization during active connectivity, while deploying Naive Bayes directly into local firmware as an instantaneous fail-safe backup during network outages.

The practical application of this data-driven framework transforms passive livestock observation into a proactive climate control ecosystem, mitigating heat stress risks in smallholder goat facilities. While the dataset utilized noise-injected simulations derived directly from the 4.2%–8.7% physical sensor error margins of DHT22 and MQ-135 calibrations, the Hardware-in-the-Loop validation achieved a 99.2% telemetry packet delivery success rate, confirming the system's operational deployment readiness.

To overcome the operational limitation of laboratory-scale simulation, future research must prioritize long-term, field-based in-vivo physical deployments inside active goat facilities (such as the Lembu Jaya barn) to evaluate system performance against live biological interactions, animal movement disturbances, and unpredictable seasonal shifts. Additionally, technical developments should expand the predictive feature matrix by integrating direct livestock physiological biomarkers—such as continuous respiratory rate variations or infrared skin temperature fluctuations—to establish a fully holistic, highly responsive smart-farming automation ecosystem.

ACKNOWLEDGMENT

The authors would like to express their highest gratitude to the Faculty of Vocational Studies, Universitas Brawijaya (Fakultas Vokasi Universitas Brawijaya) for providing the financial support and research funding that contributed to the success of this study.

REFERENCES

- Azmi, F. et al. (2025). Intelligent Actuator Control in Smart Agriculture through Machine Learning and Sensor Data Integration. *Zero : Jurnal Sains, Matematika, Dan Terapan*, 9(1), 150–161. <https://doi.org/10.30829/zero.v9i1.24421>
- Bafdal, N., Ardiansah, I., & Asmara, S. (2022). Application of Internet of Things (IoT) on Microclimate Monitoring System in The ALG Unpad Greenhouse Based on Raspberry Pi. *Jurnal Teknik Pertanian Lampung*, 11(3), 518–530.
- Barsagadea, A. G., & Rumaleb, A. S. (2024). INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Internet of Things Based Intelligent monitoring and Controlling of Poultry System on using Artificial Intelligence. *International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING*, 12, 456–467.
- Bhattacharyya, D., & Kundu, U. K. (2025). The Role of Industrial IoT and Machine Learning in Reshaping Predictive Maintenance Strategies. *Journal of Emerging Trends in Computer Science and Applications*, 1(1), 41–54.
- Boateng, F. L., Effah, E., & Abdel-fatao, H. (2026). *Turkish Journal of Agriculture - Food Science and Technology Comprehensive Review of Multi-Sensor IoT Applications in Modern Agriculture*. 14(5), 1387–1398.
- Brocklehurst, S., Fang, Z., & Butler, A. (2025). Real-Time Auto-Monitoring of Livestock : Quantitative Framework and Challenges. *Sensors*, 25(5871), 1–58.

- Candra, A. P. (2025). Analisis Data Menggunakan Python : Memperkenalkan Pandas dan NumPy. *Journal of Information System and Education Development*, 3(1), 11–16.
- Chundawat; et all. (2024). Intelligent IoT based Temperature and Humidity Monitoring System for Tobacco Curing Barn. *2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS)*, 419–423. <https://doi.org/10.1109/ICICNIS64247.2024.10823279>
- Fauza, A., Hidayat, S., Ikhsan, A., & Wulandari, B. (2022). Sistem Informasi Monitoring Integrasi Program pada Direktorat PMPUPU Badan POM. *Jurnal Informatika Universitas Pamulang*, 7(1), 9–20.
- Japit, S., Risyani, Y., Bombongan, C., & Selamat, T. (2025). Pemanfaatan Pandas , NumPy , Spark , dan TableauPower BI dalam Pengolahan dan Visualisasi Big Data untuk Mendukung Pengambilan Keputusan. *Jurnal Minfo Polgan*, 14(1), 452–456.
- Jauhari Lutfil, D. I. (2026). Sistem otomatisasi ventilasi, pemanas, dan pembersihan kotoran berbasis iot menggunakan esp-32 iot-based ventilation, heating, and cleaning automation system using esp-32. *Journal of Information Technology and Computer Science (INTECOMS)*, 9(2), 172–181.
- Jebari, H., Rekiek, S., & Reklaoui, K. (2025). Advancing Precision Livestock Farming : Integrating Hybrid AI , IoT , Cloud and Edge Computing for Enhanced Welfare and Efficiency. *International Journal of Advanced Computer Science and Applications*, 16(7), 302–311.
- Jerhamre, E., Johan, C., Carlberg, C., & Zoest, V. Van. (2022). Smart Agricultural Technology Exploring the susceptibility of smart farming : Identified opportunities and challenges. *Smart Agricultural Technology*, 2(August 2021), 100026. <https://doi.org/10.1016/j.atech.2021.100026>
- Kumar, V., Vaibhav, K., Kedam, N., Patel, A., Ram, T., & Rathnayake, U. (2024). Smart Agricultural Technology A comprehensive review on smart and sustainable agriculture using IoT technologies. *Smart Agricultural Technology*, 8(February), 100487. <https://doi.org/10.1016/j.atech.2024.100487>
- Mayberry, D., Ash, A., Prestwidge, D., & Herrero, M. (2018). Closing yield gaps in smallholder goat production systems in Ethiopia and India. *Livestock Science*, 214(January), 238–244. <https://doi.org/10.1016/j.livsci.2018.06.015>
- Nasihin, A., & Fajar, D. M. (2026). HEAT STRESS CHALLENGES AND MITIGATION STRATEGIES IN INDONESIAN SMALLHOLDER DAIRY FARMS : A BRIEF REVIEW. *Agrinimal Jurnal Ilmu Ternak Dan Tanaman*, 14(177), 114–123.
- Nejad, J. G., & Sung, K. (2017). Behavioral and physiological changes during heat stress in Corriedale ewes exposed to water deprivation. *Journal of Animal Science and Technology*, 59(13), 1–6. <https://doi.org/10.1186/s40781-017-0140-x>
- Perumal, P. (2025). Role of housing microclimate in shaping physiological , hormonal , and productive traits in goats under tropical stress conditions. *Tropical Animal Health and Production* (2025), 57(522).
- Purba, M. (2025). Klasifikasi Jenis Daun Tanaman Tropis Menggunakan Model ResNet50 Berbasis Transfer Learning dengan Dataset Tropical Plant Leaf. *JCOSIS (Journal Computer Science and Information Syetem)*, 2(2), 51–57.
- Renata Pilarczyk; et al. (2026). The Influence of Environmental Conditions and Husbandry Practices on Goat Welfare. *Animals*, 16(838), 1–30.
- Sahapudeen, F. F., Reyaa, K., Rubika, R., & Subhadharani, J. (2024). Live Stock Shelter Management System using IoT and ML. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, 4(7), 48175. <https://doi.org/10.48175/IJARSCT-17816>
- Sathees, S., Priyanka, M., Renusree, T., Samyuktha, S. K., & Sandhiya, L. (2025). LIVE STOCK MONITORING AND MANAGEMENT SYSTEM USING A RENTABLE IOT FRAMEWORK FOR RURAL CATTLEOWNERS. *International Research Journal of Modernization in Engineering Technology and Science*, 07(05), 9048–9056.
- Sejian, V., Silpa, M. V., Nair, M. R. R., Devaraj, C., Krishnan, G., Bagath, M., Chauhan, S. S., Suganthi, R. U., Fonseca, V. F. C., König, S., Gaughan, J. B., Dunshea, F. R., & Bhatta, R. (2021). Heat Stress and Goat Welfare: Adaptation and Production Considerations. *Animals*, 11(1021), 1–24.
- Setiawan, A. A., Erwanto, E., Hartono, M., & Qisthon, A. (2021). Pengaruh Manipulasi Iklim Kandang Melalui Pengkabutan Terhadap Respon Fisiologis Dan Ketahanan Panas Kambing Sapera Dan Peranakan Ettawa. *Jurnal Riset Dan Inovasi Peternakan (Journal of Research and Innovation of Animals)*, 5(1), 64–69. <https://doi.org/10.23960/jrip.2021.5.1.64-69>
- Sharma, S., Sharma, C., Asenso, E., & Sharma, K. (2023). Research Constituents and Trends in Smart Farming : An Analytical Retrospection from the Lens of Text Mining. *Hindawi Journal of Sensors*, 2023, 21. <https://doi.org/10.1155/2023/6916213>
- Silva, W. C. da. (2023). Evaluation of the temperature and humidity index to support the implementation of a rearing system for ruminants in the Western Amazon. *Frontiers in Veterinary Science*.
- Syarif, M. (2022). WATERFALL SEBAGAI MODEL PENGEMBANGAN SISTEM. *Jurnal Teknologi Informasi*, 6(1), 44–52.

- Terence, S., Immaculate, J., & Raj, A. (2024). Systematic Review on Internet of Things in Smart Livestock Management Systems. *Sustainability*, *16*(4073), 1–37.
- Zhang, L., Hang, Z., & Hu, H. (2024). Toward Large-scale Application of a Bio-inspired MEMS Optical Hair Flow Sensor Array. *2024 IEEE SENSORS*, *1*(d), 1–4.
<https://doi.org/10.1109/SENSORS60989.2024.10784825>