



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Integrated AI-Based Cybersecurity Risk Mitigation Framework For Precision Agriculture

Shubham Kumar^{1*}, Mohammad Faisal²

^{1,2} Department of Computer Application, Integral University Lucknow, UP India, E mail: kshubham@iul.ac.in, shubhamsarkar07@gmail.com

Abstract

Precision agriculture (PA) increasingly relies on interconnected IoT devices, communication networks, cloud services, and intelligent systems, thereby expanding the agricultural cybersecurity attack surface. Conventional intrusion-detection approaches mainly identify malicious traffic such as normal (0) or malicious (1) but provide limited support for explainable risk prioritization and mitigation decision-making. This study proposed a CSRM framework that establishes an end-to-end computational way from agricultural network-flow analysis. The framework integrates data processing, feature engineering, feature selection, RF-based binary threat detection, attack-probability, SHAP-based explain-ability, quantitative risk scoring, risk-level classification, and deterministic mitigation mapping. The framework has implemented and evaluated through the Farm-Flow dataset containing 1,309,887 network-flow observations, splitting 80:20 Farm-flow datasets for training-testing. Further, using 40 selected features and 200 RF trees, the proposed detector achieved accuracy (AR) (99.9282%), precision (PS) (99.9434%), recall (RC) (99.8811%), F1-score (FS) (99.9122%), ROC-AUC (0.999519), Average Precision (AP) (0.999982), and MCC (0.958845). The false-negative rate (FNR) has 0.1189%, with 256,181 attacks correctly detected from 261,978 test observations. SHAP analysis identified packet volume, flow duration, temporal characteristics, and directional traffic relationships as influential predictive attributes. The risk-assessment layer classified observations into various levels and mapped these categories level to corresponding mitigation priorities. The combined classification-probability inference throughput reached 118,538 records/s.

Keywords: Precision Agriculture, Agricultural IoT, Cybersecurity, Intrusion Detection, Random Forest, XAI, SHAP, Cybersecurity Risk Assessment, Risk Mitigation, Smart Agriculture.

1. Introduction

The integration of the IoT, wireless communication, cloud computing, edge intelligence, and AI has transformed conventional agriculture into a highly interconnected and data-driven environment. PA increasingly relies on sensors, automated machinery, communication networks, unmanned systems, and intelligent analytics to monitor crop, environmental, and operational conditions and support timely decision-making. However, this increasing connectivity also expands the cyberattack surface of agricultural systems and introduces risks to data integrity, service availability, and operational continuity [1–2].

PA settings differ from conventional IT infrastructures because they combine resource-constrained IoT devices, heterogeneous communication protocols, sensors, actuators, gateways, cloud services, and physical processes. Consequently, cyberattacks such as denial-of-service, spoofing, botnet activities, unauthorized access, and malware have affected both digital resources and agricultural operations [1][3]. Recent studies have focused on growing significant of intrusion detection, cybersecurity risk assessment (CSRA), monitoring, and protection approaches in smart agriculture (SA) [2]. The emergence of agriculture-specific datasets, especially the Farm-Flow (FF) dataset, has further enabled the development and evaluation of ML-based intrusion detection methods for agricultural IoT networks [4]. Nevertheless, detecting an attack does not by itself indicate its risk level or determine which security response have received priority.

Cybersecurity in PA have complicated by the heterogeneity and resource constraints of agricultural IoT settings. Devices have differed considerably in computational capacity, communication protocols, software configurations, and security capabilities, making the deployment of uniform security controls difficult [1][3]. In addition, agricultural networks have exhibit diverse 1 behavior, including flooding, scanning, spoofing,

botnet-based activities, and protocol-related attacks. The imbalanced distribution between normal (0) and malicious (1) traffic further makes reliable detection challenging.

More significantly, conventional IDSs mainly focus on determining whether a network events have 0 or 1. For operational security, however, detection have complemented by information about the likelihood and severity of a threat and by evidence that provides security analysts understand the method decision. A practical cybersecurity framework has therefore established a clear pathway from attack detection to CSRA and mitigation rather than terminating at classification.

Although ML and DL techniques have demonstrated promising outcomes for agricultural IDSs, existing methods largely concentrate on attack detection or classification [4]. Limited attention has been given to integrating detection outcomes with explainable risk prioritization and subsequent mitigation decisions within a single PA cybersecurity framework.

Three major gaps are identified. First, there is limited integration between AI-based threat detection and quantitative cybersecurity risk assessment. Second, detection models may provide insufficient information about the network-flow characteristics contributing to individual predictions. Third, the transition from detected threats to prioritized mitigation actions remains insufficiently addressed in an integrated agricultural cybersecurity workflow. These limitations motivate the development of a framework that connects threat detection, attack probability, explainability, risk quantification, risk classification, and mitigation.

To address these gaps, this study proposed CSRMF for PA. The framework has implemented and evaluated using the Farm-Flow dataset [4]. Its objective has extended conventional intrusion detection by transforming detection outcomes into interpretable and prioritized cybersecurity risk information.

The primary research contributions of this study:

- A RF-based ML pipeline is developed to detect malicious (0) network activity within the Farm-Flow agricultural IoT environment
- SHAP-based analysis is employed to identify influential network-flow features and provide interpretable evidence for the model's threat predictions.
- A normalized cybersecurity risk score is formulated by integrating attack probability, threat severity, and SHAP-derived evidence to quantify the potential risk associated with detected network events.
- Detected events are categorized into predefined risk levels, and corresponding mitigation priorities are assigned to support systematic cybersecurity decision-making.
- Data processing, feature engineering, feature selection, AI-based threat detection, attack-probability estimation, explainability, risk assessment, mitigation, and security response are integrated into a unified and reproducible cybersecurity framework.

Unlike detection-only approaches, the proposed CSRMF framework aims to bridge the gap between ML-based IDS and practical cybersecurity risk prioritization.

The remainder of this paper is structured as follows. Section 2 presents the related work and research gap, while Section 3 describes the proposed CSRMF. Section 4 details the dataset, preprocessing, experimental protocol, environment, and evaluation metrics. Section 5 experimental results. Section 6 provides comparative analysis and discussion. Section 7 conclusion of the study and future work.

2. Related Work

Recent research on PA cybersecurity has focused mainly on threat identification, AI-based intrusion detection, cybersecurity risk assessment (CSRA), and incident response. However, these components are often investigated independently. This part reviews the major research directions and identifies the gap addressed by the proposed integrated framework.

2.1 Cybersecurity in Precision Agriculture

The integration of IoT sensors, actuators, wireless communication, cloud computing, and intelligent systems has increased the efficiency of PA while creating new cybersecurity vulnerabilities. Agricultural environments are exposed to threats including spoofing, DoS, botnet activity, scanning, malware, and unauthorized access [9-10]. The heterogeneous and resource-constrained nature of agricultural IoT devices further complicates the deployment of conventional security approaches [11]. FF agriculture dataset has developed specifically to support IDSs research in PA using network-flow characteristics [12]. It provides agriculture-specific 0 and 1 traffic and has subsequently become a significant benchmark for evaluating AI-based agricultural IDSs.

Nevertheless, detecting an attack alone does not determine its cybersecurity priority or the mitigation action required.

2.2 AI-Based Intrusion Detection

ML and DL have been widely investigated for detecting 1 behavior in agricultural IoT networks [9][12]. Farm-Flow-based studies have demonstrated the effectiveness of both conventional ML and advanced DL models for attack detection and classification [12-14]. These methods evaluated performance through matrices including accuracy, precision, recall, F1-score, and ROC-AUC. Despite their effectiveness, most existing IDS approaches mainly outcomes whether network traffic is 1. They provide comparatively limited support for determining how serious the detected threat is, which event should receive priority, and what mitigation should be applied. Therefore, detection performances have insufficient for completed cybersecurity risk management.

2.3 Cybersecurity Risk Assessment

CSRA have provided an approach for prioritized of security events according to potential impact. In AIoT, data-driven risk assessment have identified to reduced subjectivity associated with expert-based approaches [15]. Recent PA studies also emphasize risk assessment as an significant factors of comprehensive cybersecurity management [10][16]. However, risk assessment has often treated separately from AI-based IDSs. A prediction probability generated by an IDS has not necessarily represented the completed cybersecurity risk of an event. Effective risk assessment has therefore considered multiple forms of evidence, including the probability of attack, threat severity, and evidence supporting the method decision.

2.4 Risk-Based Cybersecurity Mitigation

Cybersecurity mitigation (CSM) provides a selecting appropriated responses after identified and assessment a threat. Existing agriculture security and incident security-response frameworks address activities such as threat analysis, containment, recovery, and forensic investigation [10][17]. These methods demonstrate the significant of moving beyond detection toward structured response. However, a clear computational linkage between AI-based detection, quantified risk, and attack-specific mitigation remains limited. Consequently, the output of an IDS has not always directly transformed into an actionable security response. A risk-based mitigation approach has established a systematic way from detected threat to risk level and corresponding mitigation strategy.

2.5 Research Gap

The related work focus on research gaps. An existing agricultural IDS studies mainly emphasize attack detection and classification, rather than cybersecurity risk prioritization. AI-based detection and risk assessment are commonly treated as separate processes, limiting their integration into a unified security workflow. Although explainability techniques are increasingly used to interpret AI predictions, their evidence is rarely incorporated directly into cybersecurity risk scoring. Existing farming security frameworks provide mitigation strategies but generally lack an explicit computational relationship between detected attack, risk level, and attack-specific mitigation.

To address these gaps and limitations, and proposed an existing integrated frameworks for SA and PA:

Table 1: Comparative of Existing Cybersecurity and AI-based framework in SA and PA

Ref.	Research Focus	Technique	Dataset / Context	Finding	Limitation
[9]	Agricultural IoT intrusion detection	ML-based IDS	Agriculture 4.0	Demonstrated the applicability of ML for agricultural intrusion detection	Focuses primarily on attack detection
[10]	Smart-agriculture cybersecurity	SLR	58 studies	Identified major threats, IDS approaches, frameworks, and security challenges	Limited integration of detection, risk assessment, and mitigation
[11]	Smart-agriculture cyber threats	Threat analysis and security review	Smart Agriculture	Categorized threats, vulnerabilities, and protection approaches	Does not provide an integrated AI-driven mitigation pipeline

[12]	Agriculture-specific intrusion detection	ML-based network-flow analysis	Farm-Flow	Established an agriculture-specific benchmark for intrusion detection	Primarily evaluates intrusion detection rather than risk-based mitigation
[13]	Agricultural IoT intrusion detection	Lightweight Edge Deep Learning	Multiple IoT IDS datasets including Farm-Flow	Investigated computationally efficient deep-learning-based detection	Detection-oriented; limited risk prioritization and mitigation
[14]	Agricultural IoT intrusion detection	FICNet / Deep Learning	Farm-Flow	Achieved high binary and multiclass intrusion-detection performance	Does not establish an integrated risk-to-mitigation mechanism
[15]	Agricultural IoT risk assessment	RF-based risk assessment	Agricultural IoT	Applied ML for data-driven security-risk assessment	Risk assessment is not directly integrated with AI-based IDS outputs
[17]	Agricultural cybersecurity response	Digital Forensics and Incident Response (DFIR)	IoT-based SA	Provides structured incident analysis and response	Limited integration with quantitative AI-generated risk scores
Proposed Method	Integrated cybersecurity risk mitigation framework (CSRMF)	RF + SHAP + Probability + Severity + Risk Scoring + Risk-Level Classification + Attack-Specific Mitigation	Farm-Flow	Connects threat detection with explainable risk prioritization and mitigation	Single-dataset evaluation; broader generalization is future work

In Table 1, the novelty of the present research has not rest simply on applying another classifier to the FF dataset. Rather, these lies in establishing CSRMF pipeline in which threat detection serves as the starting point for explainable risk quantification and attack-specific mitigation. This distinction positions the proposed work beyond conventional agricultural IDSs studies and focused addresses the identified detection-to-mitigation gap.

3. Proposed Cybersecurity Risk Mitigation Framework (CSRMF)

The proposed CSRMF integrates data processing, feature engineering, feature selection, RF-based threat detection, attack-probability estimation, SHAP-based explainability, risk assessment, risk classification, and mitigation mapping for precision-agriculture cybersecurity. Using the Farm-Flow dataset [12], the framework transforms network-flow data into an interpretable and risk-aware cybersecurity decision (CSD) [20].

3.1 CSRM Framework Architecture

The proposed CSRMF consists of interconnected stages that transform raw Farm-Flow network traffic into an interpretable CSD. The first step performs data preprocessing and feature preparation. The processed features are then supplied to the RF model for threat detection. Following detection, the model-generated attacks probability has obtained and SHAP analysis have performed to provide feature-level evidence for the prediction [20].

The probability, severity, and SHAP-based evidence are subsequently incorporated into the risk-assessment stage. The resulting risk score have converted into an operational risk level. Finally, the risk levels have mapped to a predefined mitigation action.

The CSRMF architecture:

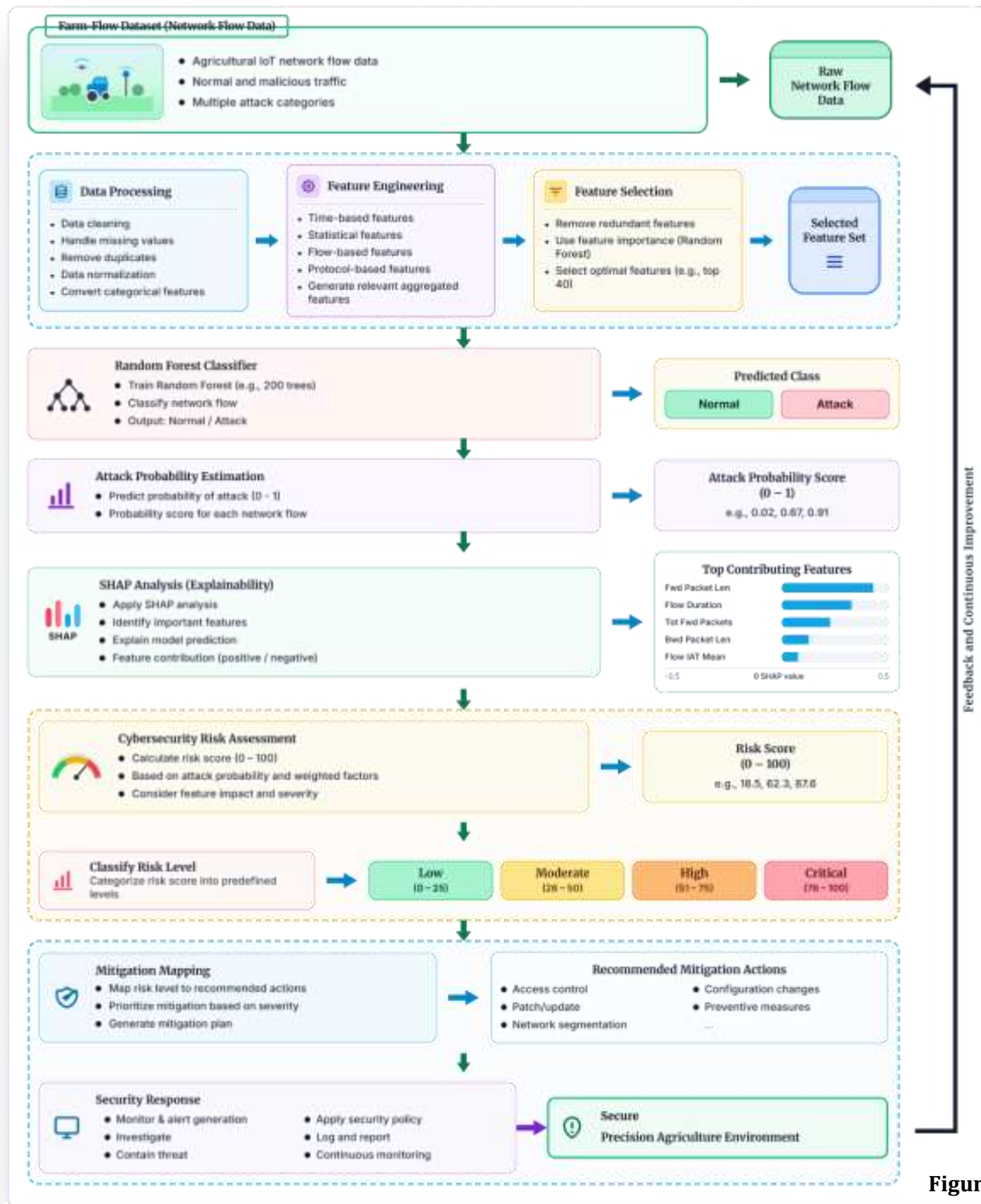


Figure 1:

Working Flowchart of Cybersecurity Risk Mitigation Framework (CSRMF)

CSRMF architecture extends conventional intrusion detection by connecting attack identification with risk prioritization and mitigation selection.

3.2 Farm-Flow Data Processing

The Farm-Flow agriculture dataset contains network-flow observations generated within an agricultural IoT setting and includes both normal and 1 traffic [12]. Before applying the AI-based detection model, the dataset has subjected to systematic preprocessing.

The data-processing steps including data cleaning, missing-value handling, duplicate checking, encoding of categorical attributes where required, feature engineering, and feature selection. These operations have performed to obtain a consistent and informative representation of the network-flow data. After preprocessing, the input data have represented as:

$$X = \{x_1, x_2, \dots, x_n\} \quad (1)$$

Where X represents the set of selected network-flow features and x_i represents an individual feature. Equation (1) indicates that the original Farm-Flow records have transformed into a set of features that have provided to the ML model. Because RF has a tree-based ensemble algorithm, feature normalization has not inherently required for its classification process [19]. Therefore, unnecessary feature scaling has avoided.

3.3 AI-Based Threat Detection

These layer identifies 1 network activity from the selected Farm-Flow network-flow features. The proposed framework employs a RF classifier to distinguish between normal and attack traffic and to generate the corresponding attack probability. The resulting predictions and probability estimates are subsequently used by the risk-assessment and mitigation layers.

3.3.1 Random Forest (RF)-Based Threat Detection

RF has applied as mainly AI-based threat-detection model. It is an ensemble ML method that combines various decision trees (DT) to obtain a classification result [19].

The processed Farm-Flow features have provided to the trained RF model. The predicted traffic or attack class has represented as:

$$\hat{y} = \text{RF}(X) \quad (2)$$

Where X represents the processed network-flow features, RF represents the trained RF classifier, and $\hat{y}$ represents the predicted class.

3.3.2 Attack-Probability Estimation

After identifying the traffic or attack class, the RF model provides the probability associated with the predicted class. For an attack class a , the probability has expressed as:

$$(P_a) = P(y = a | X) \quad (3)$$

Where P_a represents the estimated probability of attack class a , $y = a$ represents the corresponding class, and X represents the observed network-flow features. If $P_a = 0.95$, the model assigns a 95% probability to the identified attack class.

Therefore, Equation (3) provides a quantitative indication of how likely the detected attack is, which is subsequently used in the risk-assessment step.

3.4 Cybersecurity Risk Assessment (CSRA)

The risk-assessment step converts the outputs of threat detection into a quantitative representation of cybersecurity risk. The proposed framework considers attack probability, threat severity, and SHAP-based evidence.

3.4.1 SHAP-Based Explainability

SHAP have used to interpreted the predictions generated by the RF model [20]. It identifies the contribution of individual features to the model prediction.

The SHAP-based evidence associated with a prediction have represented as:

$$E = \text{SHAP}(X, \hat{y}) \quad (4)$$

Where E represents the SHAP-based evidence, X represents the input features, and $\hat{y}$ represents the predicted class. The purpose of this steps have to determine which network-flow features provide evidence for the detected threat. This improves the interpretability of the subsequent risk decision.

3.4.2 Cybersecurity Risk Score

The proposed framework combines attack probability, threat severity, and SHAP-based evidence to calculate the cybersecurity risk score:

$$R = w_p P_a w_s S + w_e E \quad (5)$$

where:

- R represents the cybersecurity risk score;
- P_a represents attack probability;
- S represents threat severity;
- E represents SHAP-based evidence;

- w_p represents the weight assigned to attack probability;
- w_s represents the weight assigned to severity; and
- w_e represent the weight assigned to SHAP evidence.

The weighting factors satisfy:

$$w_p + w_s + w_e = 1 \quad (6)$$

Equation (5) means that the cybersecurity risk is determined by considering how likely the attack is, how severe it is, and how strongly the model evidence supports the detected threat.

The resulting risk score have represented on a 0–100 scale, where a higher value indicates greater cybersecurity risk.

3.4.3 Risk-Level Classification

After calculating the risk score, the numerical value has converted into an operational risk level. The proposed framework uses four risk categories:

$$L = \begin{cases} \text{Low}, & 0 \leq R < 25 \\ \text{Moderate}, & 25 \leq R < 50 \\ \text{High}, & 50 \leq R < 75 \\ \text{Critical}, & 75 \leq R \leq 100 \end{cases} \quad (7)$$

Equation (7) converts the numerical risk score into an understandable security condition that have used to determine the required response.

3.5 Risk-Based Cybersecurity Mitigation Strategy

The proposed framework uses a deterministic risk-to-action mapping to associate each risk level with a corresponding mitigation response. The approach provides a transparent relationship between the estimated cybersecurity risk and the recommended security action, presented in Table 2.

Table 2: Deterministic Risk-to-Mitigation Mapping

Risk Level	Risk Score	Recommended Mitigation	Security Response
Low	0–24.99	Monitor	Continue monitoring the identified traffic
Moderate	25–49.99	Alert / Rate Limiting	Generate an alert and restrict suspicious traffic
High	50–74.99	Block Suspicious Traffic	Block or restrict suspicious communication
Critical	75–100	Isolate Affected Node + Alert	Isolate the affected agricultural IoT node and generate a high-priority alert

The mitigation action is selected according to the identified risk level and have represented as:

$$M = g(L) \quad (8)$$

Where M represents the selected mitigation action, L represents the risk level, and g represents the predefined risk-to-mitigation mapping.

The mapping has defined as:

$$M = \begin{cases} \text{Monitor}, & L = \text{Low} \\ \text{Alert/Rate Limiting}, & L = \text{Moderate} \\ \text{Block Suspicious Traffic}, & L = \text{High} \\ \text{Isolate Affected Node + Alert}, & L = \text{Critical} \end{cases} \quad (9)$$

The proposed framework consequently establishes a complete detection-to-risk-to-mitigation pathway. Farm-Flow network traffic have first transformed into suitable ML features, after which RF identifies the corresponding traffic or attack class. The attack probability and SHAP-based evidences have then considered together with threat severity to obtain the cybersecurity risk score. The scores have converted into a defined risk level, and the resulting level determines the recommended mitigation response.

Thus, the proposed CSRMF provided an integrated and interpretable method for transforming agricultural network-flow observations into prioritized cybersecurity decisions without introducing an additional learning or optimization component for mitigation selection.

4. Experimental Setup and Evaluation Methodology

This section described the Farm-Flow agriculture dataset, preprocessing, training and testing protocol, experimental environment, and evaluation metrics used to assess the proposed CSRMF. The experiments employ RF for binary threat detection, followed by risk assessment, classification, mitigation prioritization, and computational evaluation. A fixed random seed of 42 have used to ensured reproducibility.

4.1 Farm-Flow Agriculture Dataset

The Farm-Flow agriculture dataset have used as the mainly benchmark because it represented network-flow traffic from a PA IoT setting [12]. The dataset contains 1,309,887 observations and 101 columns, including 0 and 1 traffic. In this study, network identifiers and label-related fields have excluded from the predictor set, while `is_attack` has used as the binary target, where 0 denotes Normal and 1 denotes Attack. After preprocessing and feature selection, 40 features have retained for model development. The dataset exhibits substantial class imbalance, with 27,458 normal (2.10%) and 1,282,429 attack (97.90%) observations, presented in Table 3. This imbalance has addressed using the RF `class_weight="balanced"` setting without synthetic oversampling.

Table 3: Farm-Flow Dataset Characteristics

Characteristic	Description
Dataset	Farm-Flow
Application domain	Smart agriculture / AG-IoT cybersecurity
Total instances	1,309,887
Original columns	101
Original processed binary setting	29 features reported in the original study
Features entering current ML pipeline	90
Final selected features	40
Target variable	<code>is_attack</code>
Classification formulation	Binary
Class 0	Normal
Class 1	Attack
Normal records	27,458 (2.10%)
Attack records	1,282,429 (97.90%)
Attack categories represented	ARP_Spoofing, BotNet_DDoS, HTTP_Flood, ICMP_Flood, MQTT_Flood, Port_Scanning, TCP_Flood, UDP_Flood

The current implementation therefore uses the Farm-Flow dataset as a single-dataset experimental benchmark, without introducing an external dataset into the evaluation pipeline. The final project configuration explicitly identifies Farm-Flow as the dataset and RF as the model used in the framework.

4.2 Data Preprocessing

Data preprocessing involved data cleaning, feature engineering, feature reduction, and missing-value treatment. Missing-value representations have standardized as NaN, and the binary target `is_attack` have validated. Non-predictive identifiers and label-related fields have excluded, followed by removal of constant and near-constant features. The initial feature set contained 90 features (83 numeric and 7 categorical).

Feature engineering generated flow-level characteristics related to packet and byte ratios, traffic rates, payloads, TCP flags, flow duration, header sizes, inter-arrival times, subflows, bulk transfers, TCP windows, and active/idle behaviour. Safe division have applied to avoid numerical instability, with infinite values subsequently treated as missing.

Feature selection used a 0.95 correlation threshold to reduce redundancy, followed by RF-based feature ranking on a stratified sample of 200,000 observations. A 100-tree RF model with balanced class weighting have used for ranking, and the top 40 features have retained. Median imputations have fitted on the training data and applied to the test data, with final quality checks confirming no unresolved missing or infinite values.

4.3 Training and Testing Protocol

A stratified 80:20 train-test split form-flow dataset have applied to preserved the class distribution, producing 1,047,909 training and 261,978 testing observations. The detection model used a 200-tree RF classifier with Gini criterion, balanced class weighting, square-root feature selection, and `random_state=42`.

The models have trained exclusively on the training subset, while the independent test set has used to generate binary predictions and attack probabilities. The predicted probabilities have subsequently used for risk assessment, risk classification, and mitigation mapping. No external validation dataset has used. The trained model, test data, feature information, SHAP outputs, probabilities, and risk results have retained to support reproducibility.

4.4 Experimental Environment

The experiments have implemented in Python language using the scientific computing and ML ecosystem. The completed reproducibility record reports, and execution setting has a 64-bit Windows 10 system based on an AMD64 architecture. The RF implementation has used parallel processing through `n_jobs=-1`. The final CSRMF contains 200 trees and operates on 40 selected features, as a shown in Table 4. The stored computational-performance results additionally record model size, test-data memory consumption, prediction time, and throughput, allowing the proposed CSRMF to be evaluated not only from a predictive perspective but also from an operational-efficiency perspective.

Table 4: Experimental Configuration

Parameter	Configuration
Programming language	Python 3.13.9
Operating system	Windows 11, 64-bit
Dataset	Farm-Flow
Classification task	Binary attack detection
Final input features	40
Training/testing split	80% / 20%
Split strategy	Stratified
Random state	42
Classifier	RF
Number of trees	200
Criterion	Gini
Maximum_depth	None
Minimum dataset split	2
Minimum dataset leaf	1
Maximum_features	sqrt
Class_weighting	Balanced
Parallel jobs	-1
Scikit-learn	1.7.2
NumPy	2.3.5
Pandas	2.3.3

4.5 Assessment Metrics

The proposed CSRM framework has assessment from various complementary perspectives: binary attack detection, method discrimination and classification reliability, CSR quantification and risk-level assignment, and computational efficiency. This assessment rules have selected to assessed not only whether 1 network flows have detected accurately, but also whether the detected threats have translated into interpretable risk scores and corresponding mitigation actions. Such an integrated evaluation has especially relevant to PA settings, where cybersecurity methods have helping on-time operational decisions rather than provided detection results alone [1-2][4].

4.5.1. Binary Attack-Detection Metrics

The mainly detection task has formulated as a binary classification problem through `is_attack` target, where 0 and 1 traffic. The confusion matrix consists of TP, TN, FP, and FN. Based on these quantities, Accuracy (AR), Precision (PS), Recall (RC), Specificity, and F1-score (FS) have calculated.

A. Accuracy (AR)

AR measures the overall proportion of accurate classified network-flow instances. These provided a simple presented of classifier performance:

$$AR = (TP_1 + TN_2) / (TP_1 + TN_2 + FP_1 + FN_1) \quad (10)$$

Although AR provided an intuitive overall measure, these have not interpreted in isolation when class distribution has unequal. Therefore, additional class-sensitive metrics have considered in the present study [21].

B. Precision (PS)

PS quantified the reliability of attack data predicted by measuring the proportion of predicted attack flows that have actually attacks:

$$PS = TP_1 / (TP_1 + FP_1) \quad (11)$$

A high PS outcomes shown that the framework generates relatively few false alarms when identifying 1 traffic. This characteristic has especially significant in PA settings because excessive false alerts have increase the operational burden on farm-network administrators.

C. Recall (Sensitivity)

Recall has used to TPR that measures the proportion of actual FF data attacks successfully detected by the classifier:

$$Recall = TP_1 / (TP_1 + FN_2) \quad (12)$$

Recall has especially significant for cybersecurity applications because FN_2 shown 1 flows that remain undetected. Consequently, a high recall outcomes shown stronger capability to identify actual cyber threats.

D. Specificity (SF)

SF or TNR that assessment the ability of the classifier to correctly recognized legitimate network traffic:

$$SF = TN_2 / (TN_2 + FP_1) \quad (13)$$

High SF indicates that the framework has distinguish benign farm-network communication from 1 activity while limiting FP_1 alerts.

E. F1-Score (FS)

Because PS and RC shown different aspects of attack-detection quality, their harmonic mean has used to obtain a balanced measure:

$$FS = 2 \times (PS \times RC) / (PS + RC) \quad (14)$$

The FC has especially useful when the cost of FP and FN decisions have both be considered and when the class distribution has influence AR.

4.5.2. ROC-AUC and Average Precision (AP)

ROC-AUC has used to evaluated the ability of the RF classifier to discriminate between 0 and 1 traffic over different classification thresholds. ROC-AUC summarizes the relationship between the TPR and FPR across the complete threshold range. A value close to 1 indicates strong discriminatory capability.

$$TPR = TP / (TP + FN) \quad (15)$$

$$FPR = FP / (FP + TN) \quad (16)$$

The ROC-AUC provided an aggregate measured of the classifier's ability to distinguish between 0 and 1 traffic across different decision thresholds. A value closer to 1 indicated stronger discriminatory performance.

Because the Farm-Flow agriculture dataset contains a highly imbalanced distribution between 0 and 1 observations, AP has also considered. AP summarizes the precision-recall (PR) relationship across different probability thresholds and has expressed as:

$$AP = \sum_n (R_n - R_{n-1}) / P_n \quad (17)$$

Where P_n and R_n denote PR and recall at the n^{th} operating point, respectively. Therefore, the combined use of ROC-AUC and AP provides a more comprehensive evaluation of attack-class discrimination than AR alone.

4.5.3. Explainability and Feature-Significant Assessment

SHAP have employed to interpreted the RF predictions and identify the network-flow characteristics contributing most strongly to attack classification. SHAP has based on additive feature attribution and provides a theoretically grounded mechanism for explaining individual model predictions [19]. The global feature-importance analysis has obtained using the mean absolute SHAP value:

$$f(x) = \phi_0 + \sum_{i=1}^m \phi_i \quad (18)$$

$$I_i = 1/N \sum_{j=1}^N |\phi_{ji}| \quad (19)$$

Where $f(x)$ is the model output, ϕ_0 denotes the expected model output, m is the number of input features, and ϕ_i denotes the contribution of feature i .

For global feature-significant analysis, the mean absolute SHAP contribution has calculated as:

$$I_i = 1/N \sum_{j=1}^N |\phi_{ji}| \quad (20)$$

Where I_i denotes the global importance of feature i , N represents the number of evaluated observations, and ϕ_{ji} denotes the SHAP contribution of feature i for observation j .

The normalized contribution of each feature has expressed as:

$$I_i(\%) = \frac{I_i}{\sum_{k=1}^m I_k} \times 100 \quad (21)$$

Where m is the total number of selected features. The SHAP results therefore provide an interpretable representation of the factors contributing to the model's attack-detection decisions.

4.5.4 Cybersecurity Risk Assessment Metrics

Following attack detection and explainability analysis, each test observation has assigned a cybersecurity risk score. The proposed risk model integrates three evidence sources: attack probability, attack severity, and SHAP-derived evidence strength. The risk score has defined as:

$$R = w_p P + w_s S + w_e E \quad (22)$$

Where P denotes the attack-probability component, S denotes the severity component, E denotes the normalized SHAP evidence component, and w_p , w_s , and w_e represent their respective weights.

In the implemented framework, the weights have set to $w_p = 0.60$, $w_s = 0.20$, and $w_e = 0.20$. Attack probability has normalized to a 0–100 scale. Severity has represented using four ordinal levels: Low = 25, Moderate = 50, High = 75, and Critical = 100. The SHAP evidence component has normalized using the model-derived SHAP strength rather than treating raw SHAP values as direct causal measurements.

The final risk score has mapped to four operational risk levels:

$$L(R) = \begin{cases} \text{Low}, & 0 \leq R \leq 25 \\ \text{Moderate}, & 25 < R \leq 50 \\ \text{High}, & 50 < R \leq 75 \\ \text{Critical}, & 75 < R \leq 100 \end{cases} \quad (23)$$

The principal risk-assessment indicators have the minimum, maximum, mean, and median risk scores and the distribution of observations across the four risk categories.

4.5.5 Risk-Based Mitigation Metrics

The mitigation stage converts the estimated risk level into an operational cybersecurity response. The implemented mapping is deterministic:

Table 5: Risk Level Mitigation

Risk Level	Risk-Score Range	Assigned Mitigation Action
Low	0–25	Routine Monitoring
Moderate	>25–50	Enhanced Monitoring
High	>50–75	Immediate Investigation
Critical	>75–100	Immediate Response

The mitigation analysis therefore evaluates the distribution of test observations across the four response categories and verifies the end-to-end mapping from risk assessment to mitigation recommendation. Since the present study evaluates the computational framework rather than field-deploying the countermeasures, mitigation effectiveness itself is not interpreted as experimentally validated intervention efficacy.

4.5.6 Computational Efficiency Metrics

Computational efficiency has evaluated using training time, model size, test-data memory consumption, inference time, prediction throughput, and prediction latency per record. Total inference time has calculated by combining classification and probability-prediction operations. Throughput has reported in records per second, while latency has expressed in milliseconds per record. These measures are important for PA cybersecurity because network-monitoring systems may operate under resource and response-time constraints [5-6].

5. Results and Discussion

The proposed CSRMF has evaluated using the held-out Farm-Flow test set to determine its effectiveness in binary attack detection, classification reliability, explainable threat analysis, cybersecurity risk quantification, risk-level prioritization, mitigation assignment, and computational efficiency. The experimental protocol produced 1,047,909 training observations and 261,978 independent testing observations using an 80:20 stratified split dataset with a fixed random state of 42. The final RF models have trained using 40 selected network-flow features and 200 decision trees. The results presented in this section are derived directly from the implemented experimental pipeline and its saved evaluation outputs. The binary detection formulation is consistent with the experimental design described in Section 4, where $is_attack = 0$ represents normal traffic and $is_attack = 1$ represents malicious traffic. The Farm-Flow dataset provides a large-scale agriculture-

specific benchmark for network-flow-based intrusion detection [4], while recent agricultural cybersecurity studies have demonstrated the applicability of ML and DL methods to this domain [8-9][13-14].

5.1 AI-Based Threat Detection Results

The RF classifier demonstrated strong binary attack-detection performance on the independent test set. As shown in Table 6, the method gained of AR of 99.9282%, showing that 99.9282% of the 261,978 test observations have classified correctly. PS obtained 99.9434%, demonstrating that the overwhelming majority of observations predicted as attacks have actually 1. RC have 99.8811%, showing that the classifier successfully identified nearly all 1 observations present in the test set. SF have 97.3598%, showing that the method has also able to correctly identify a substantial proportion of legitimate network traffic.

The FS of 99.9122% shown a strong balance between PS and RC. These have especially relevant for IDSs because a detection system has simultaneously reduced false alarms and avoid overlooking 1 traffic. The very small difference between PR and CR further shown that the classifier has not exhibit a substantial trade-off between these two way under the assessment decision threshold.

Table 6: Binary Attack-Detection Performance of the Proposed RF Model

Metric	Value
AR	99.8282%
PC	99.9434%
RC	99.8811%
SF	97.3598%
FC	99.9122%
ROC-AUC	0.999519
AP	0.999982
BA	98.6204%
MCC	0.958845
FPR	2.6402%
FNR	0.1189%
Overall Error Rate	0.1718%

The additional class-sensitive metrics reinforce the reliability of the detector. Balanced AR have 98.6204%, demonstrating that the high overall AR has not solely a consequence of the underlying class distribution. MCC have 0.958845, showing a strong correspondence between the predicted and actual binary classes. The overall error rate has 0.1718%.

The ROC-AUC value of 0.999519 achieved that the method maintains excellent discrimination between 0 and 1 network-flow observations across different probability thresholds. The corresponding ROC curve have presented in Figure 4. Similarly, the AP of 0.999982 showed that the method maintains very high PS over the evaluated recall range. These have especially informative for the present dataset because FF contains a highly imbalanced distribution between 0 and 1 observations. The PR relationship have shown in Figure 3.

The outcomes demonstrate that the RF method has highly effective as the first analytical component of the proposed CSRM framework. However, the primary contribution of the framework has not limited to maximizing binary classification performance. Instead, the detection output has subsequently used to generate attack probability, explainability evidence, cybersecurity risk scores, and mitigation priorities. Thus, the detection stage functions as the foundation for the subsequent risk-aware decision process.

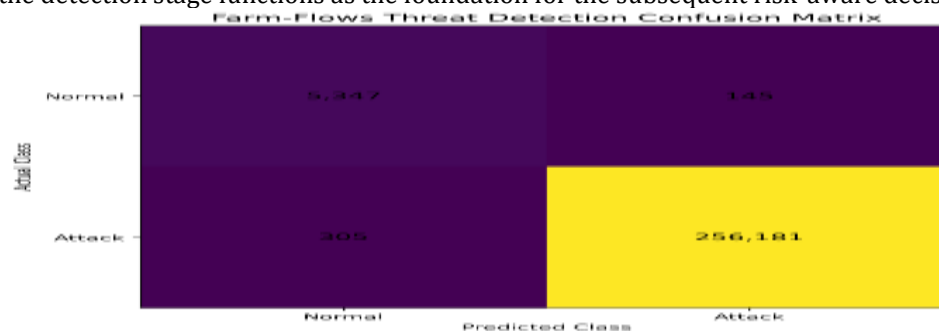


Figure 2: Confusion Matrix of RF Attack Detector

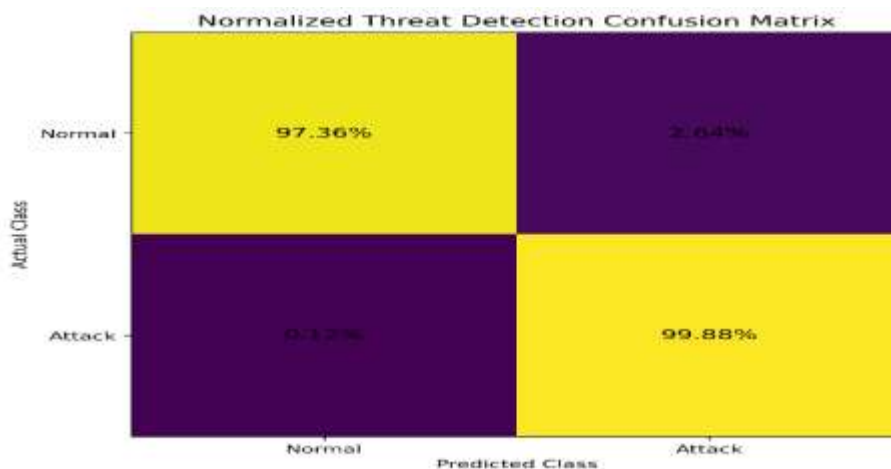


Figure 3: Normalized Confusion Matrix of RF Attack Detector

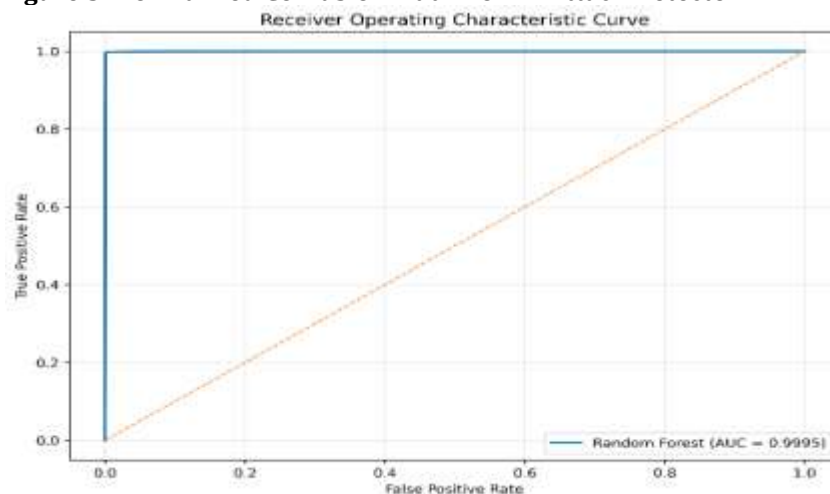


Figure 4: ROC Curve of RF Classifier

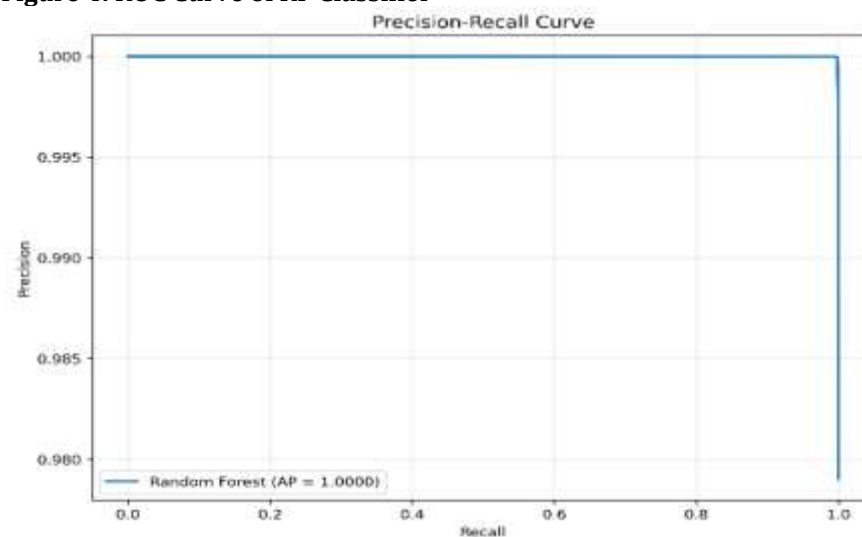


Figure 5: Precision-Recall (PR) Curve of RF Classifier

The outcomes have interpreted in the context of previous FF research. The original FF study established the dataset as an agriculture-specific network-flow benchmark for IDSs [4], while subsequent studies have investigated lightweight DL and other advanced architectures on agricultural IoT traffic [6][8][13-14]. The present outcomes have therefore consistent with the broader evidence that AI-based methods have effectively identify 1 network behavior in agricultural IoT settings. Nevertheless, the proposed CSRM framework differs in emphasis by using the detection output as an input to an integrated CSR-management process rather than terminating the analysis at classification.

5.2 Confusion Matrix (CM) and Error Analysis (EA)

The detailed CM provides additional insight into the behavior of the classifier. The test set contained 5,347 TN, 145 FP, 305 FN, and 256,181 TP, yielding a total of 261,978 observations, as ashown in Table 7.

Table 7: Confusion-Matrix-Based Error Analysis

Actual Class	Predicted Normal (0)	Predicted Attack (1)	Total
Normal	5347 (TN)	145 (FP)	5492
Attack	305 (FN)	256181 (TP)	256486
Total	5652	256326	261978

The number of TP₁ is substantially larger than the number of FN₂, outcomes in a RC of 99.8811%. This is shown that only a very small fraction of 1 flows have incorrectly classified as 0. From a cybersecurity perspective, this significant because FN₂ represent potentially undetected 1 activity.

Only 145 normal (0) observations have incorrectly classified as attacks. This produces a FPR of 2.6402%. Although FP₁ have not represented missed attacks, they have increased the workload of security operators because incorrectly flagged flows have required additional inspection. The PS of 99.9434% nevertheless shown that the number of such erroneous alerts remains small relative to the total number of predicted attack observations.

The normalized CM in Figure 3 provides the same classification behavior in proportional form and facilitates interpretation independent of the absolute number of observations. Figures 6 demonstrate that the classifier performs strongly for both classes, while maintaining especially high sensitivity to 1 traffic.

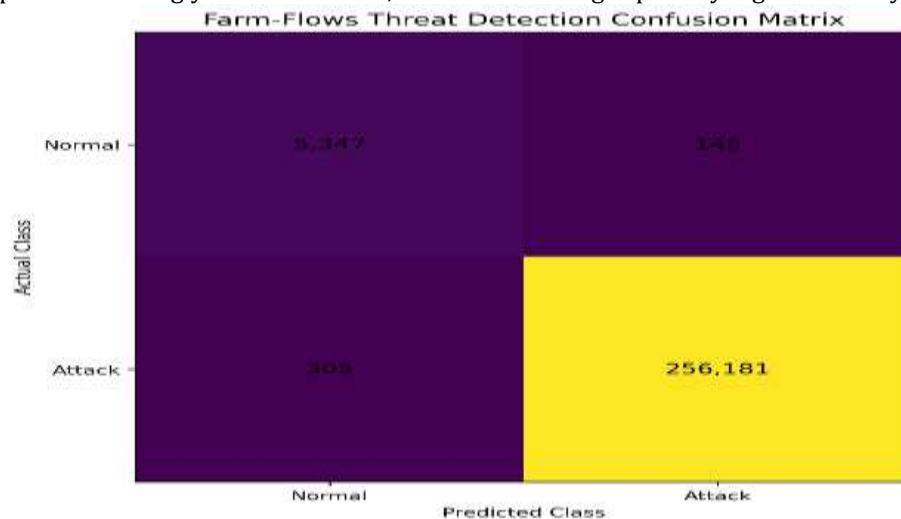


Figure 6: Confusion Matrix of the RF Attack Detector

The CM demonstrates the predominance of correctly classified attack observations and the comparatively few number of classification errors. The 305 FN₂ constitute only a small fraction of the actual attack observations, while the 145 FP₁ represent the principal source of erroneous alerts.

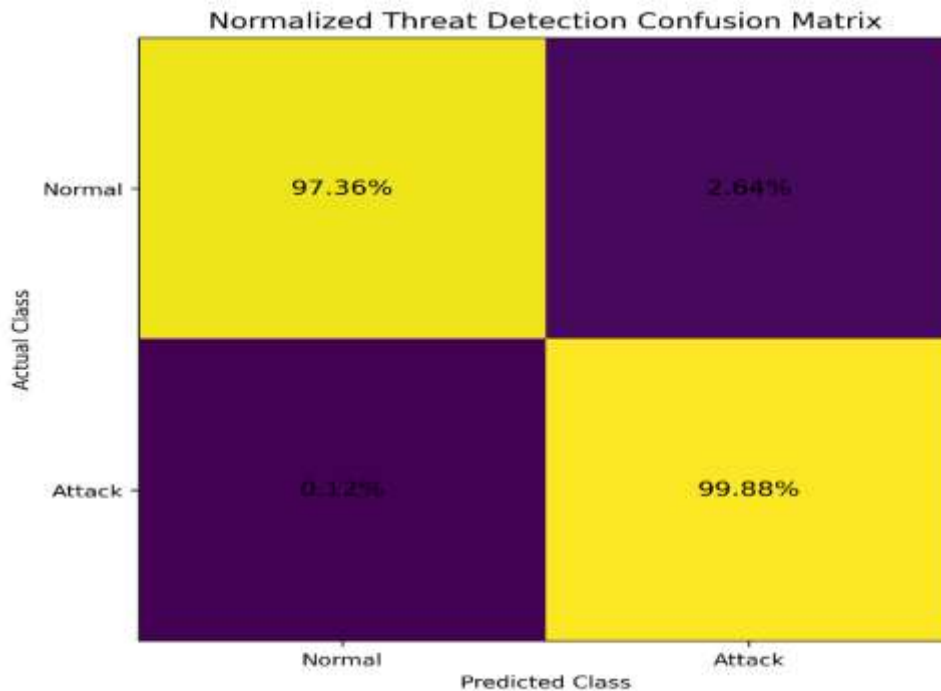


Figure 7: Normalized Confusion Matrix of RF Attack Detector

The graphical representation of normalized confirmation the high TP_1 and TN_2 recognition rates. The distinction between the two error types significant because the operational consequences of FN_2 and FP_1 have different. In farming cybersecurity settings, FP_1 have allowed 1 communication to continue unnoticed, whereas FP_1 have mainly increased monitoring and investigation workload.

The combined FPR of 2.6402% and FNR of 0.1189% shown that the implemented method places substantially greater emphasis on avoiding missed attacks under the selected operating point. This behavior has desirable for an IDSs application, although operational thresholds have adjusted according to the security requirements and acceptable alert burden of especially farming deployment.

5.3 Explainable Threat Analysis Using SHAP

To enhance the interpretability of the RF decisions, SHAP-based analysis have applied to the method predictions. The global SHAP outcomes quantify the average magnitude of each feature of contribution to the method output. This provides an interpretable view of the network-flow characteristics that have most strongly associated with the model of attack-detection decisions [19].

The outcomes shown that *orig_pkts* has the most influential feature, with a mean absolute SHAP value of 0.069406 and a relative contribution of 13.5330%. The second most influential feature has duration, with a mean absolute SHAP value of 0.051029 (9.9497%). These have followed by *active.min* (5.7221%), *fwd_iat.max* (5.4797%), and *fwd_pkts_tot* (4.4868%).

Table 8: Top Ten SHAP Features Contributing to Attack Detection

Rank	Network-Flow Feature	Mean Absolute SHAP	Relative Importance (%)
1	<i>orig_pkts</i>	0.069406	13.5330
2	<i>duration</i>	0.051029	9.9497
3	<i>active.min</i>	0.029347	5.7221
4	<i>Fwd_iat.max</i>	0.028104	5.4797
5	<i>Fwd_pkts_tot</i>	0.023011	4.4868
6	<i>Orig_resp_byte_ration</i>	0.022324	4.3528
7	<i>Fwd_pkts_per_sec</i>	0.018689	3.6440

8	Resp_orig_per_sec	0.018689	03.6440
9	Fwd_iat.min	0.018161	03.5411
10	Fwd+bwd_header_ratio	0.014785	02.8829

The dominance of packet-count and temporal-flow characteristics indicates that the methods relies substantially on traffic volume, duration, packet timing, and directional communication relationships. Features such as fwd_pkts_per_sec, fwd_iat.max, and fwd_iat.min represent traffic-rate and inter-arrival-time behavior, while ratio-based features such as orig_resp_byte_ratio and fwd_bwd_header_ratio capture directional or structural differences between communicating endpoints, presented in Table 8.

These findings have significant for cybersecurity interpretation because they provided evidence about the network-flow characteristics associated with the classifier's decisions. Rather than treating the RF as an opaque prediction mechanism, the SHAP layer identifies the attributes that most strongly contribute to the model of output. This supports the framework's objective of connecting AI-based detection with interpretable CSA.

These significant to distinguish predictive contribution from causality. A high SHAP value shown that a feature contributes strongly to the methods prediction under the evaluated data distribution. These have not established that the feature itself causes the underlying cyberattack. Therefore, SHAP evidence has used in the proposed CSRM framework as an interpretability and evidential component rather than as causal proof.

5.4 Cybersecurity Risk Assessment Results

The central objective of the proposed framework has to extended attack detection into quantitative CSRA. For each test observation, the framework combines attack probability, threat severity, and SHAP-derived evidence according to the risk formulation already defined in Section 3.

The implemented weighting configuration assigns 0.60 to attack probability, 0.20 to severity, and 0.20 to SHAP evidence. This weighting gives the model-estimated likelihood of 1 activity the largest influence on the final risk score, while severity and explainability evidence provide complementary information.

Table 9: Details of Cybersecurity Risk Scores

Risk Parameter	Result
Total test observations	261978
Minimum risk score	8.360426
Maximum risk score	83.360425
Mean risk score	81.799750
Median risk score	83.360426
Probability weight	0.60
Severity weight	0.20
SHAP evidence weight	0.20
SHAP evidence score	16.802131

The outcome CSR scores ranged from 8.360426 to 83.360426. The mean risk score has 81.799750, while the median has 83.360426. The relatively high median indicates that a large proportion of observations received high numerical risk scores under the implemented risk formulation.

This result has closely related to the attack-dominant nature of FF test population and have therefore be interpreted carefully. The risk-score distribution represents the CSRA of the assessment benchmark observations under the specified probability, severity, and SHAP weighting scheme. These have not been interpreted as an estimate of the prevalence of critical cyber incidents in practical PA networks.

Table 10: Risk-Level Distribution of the Test Observations

Risk Level	Risk-score Range	Records	Percentage (%)	Mean Risk Score	Median Risk Score
Low	0 - 25	4991	1.9051	9.7050	08.3604
Moderate	>25-50	716	0.2733	35.8153	34.9604
High	>50 - 75	270	0.1031	59.7849	57.9604
Critical	>75 - 100	256001	97.7185	83.3571	83.3604
Total	---	261978	100.0000	---	---

The risk-level analysis shows that 4,991 observations (1.9051%) have classified as Low risk, 716 observations (0.2733%) as Moderate risk, 270 observations (0.1031%) as High risk, and 256,001 observations (97.7185%) as Critical risk.

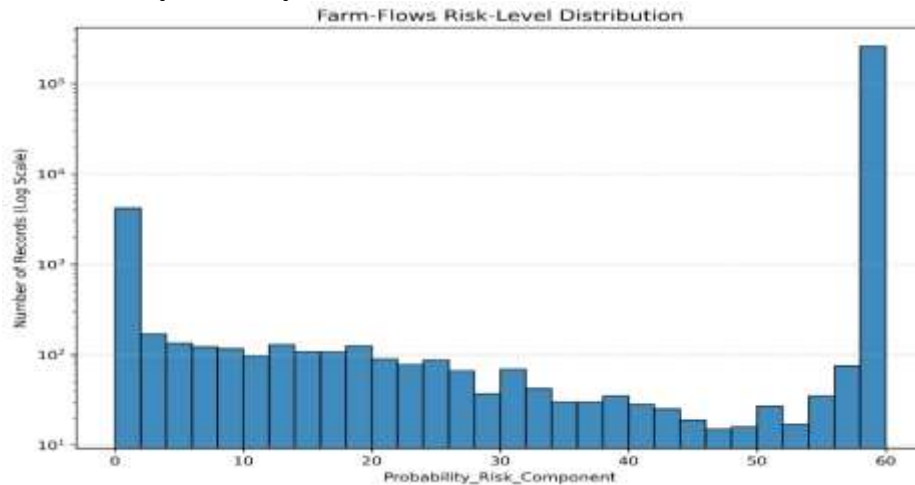


Figure 8: Risk-Level Distribution

Figure 8 showing a graphical representation of the four risk categories. The strong concentration in the Critical category has a direct consequence of the assessment benchmark population and the risk-scoring formulation. These have not been interpreted as a universal statement regarding the severity distribution of cyber events in operational agricultural systems.

The risk-level statistics further demonstrate separation between the four categories. Low-risk observations have a mean risk score of 9.7050, Moderate-risk observations have a mean of 35.8153, High-risk observations have a mean of 59.7849, and Critical-risk observations have a mean of 83.3571. The ordered progression confirms that the numerical risk scores have mapped consistently into the intended qualitative categories.

A significant property of the implementation has that all 261,978 test observations have successfully assigned a valid risk level. The implementation audit verifies that the number of classified observations equals the total number of risk records and that only the predefined Low, Moderate, High, and Critical categories occur. This ensures complete downstream coverage of the test population.

5.5 Risk-Based Mitigation Results

The mitigation stage transforms the numerical risk assessment into an operational response priority. In the implemented framework, the mapping is deterministic and follows the four risk levels defined in Section 3. Low-risk observations receive routine monitoring, Moderate-risk observations receive enhanced monitoring, High-risk observations require immediate investigation, and Critical-risk observations receive immediate response prioritization, presented in Table 11.

Table 11: Risk-Level-to-Mitigation Mapping

Risk Level	Risk Range	Mitigation Priority	Operational Response
Low	0 -25	Routine Monitoring	Continue routine monitoring; no immediate containment required
Moderate	> 25 -50	Enhanced Monitoring	Increase monitoring and perform additional traffic inspection
High	>50 – 75	Immediate Investigation	Investigation immediately and consider temporary containment
Critical	>75 – 100	Immediate Response	Initiate incident response and isolate or block suspicious communication where appropriate

The resulting mitigation distribution is presented in Table 12.

Table 12: Distribution of Risk-Based Mitigation Priorities

Risk Level	Mitigation Priority	Records	Percentage
Low	Routine Monitoring	4991	1.9051
Moderate	Enhanced Monitoring	716	0.2733
High	Immediate Investigation	270	0.1031
Critical	Immediate Response	256001	97.7185
Total	---	261978	100.0000

The mitigation results demonstrate that every evaluated observation receives a corresponding operational priority. The largest response group is Critical, containing 256,001 observations, followed by Routine Monitoring with 4,991 observations, Enhanced Monitoring with 716 observations, and Immediate Investigation with 270 observations.

A further limitation concerns attack-specific mitigation. Although Farm-Flow contains several attack categories, the implemented detection task is binary and the current mitigation output does not provide experimentally validated per-attack mitigation effectiveness. Therefore, the present results have interpreted as risk-level-based mitigation prioritization rather than validated attack-family-specific countermeasure performance.

5.5 Computational Performance

The computational evaluation examines whether the proposed detection component have processed the large Farm-Flow test set efficiently. The final RF model contains 200 DT and operates on 40 selected features. The measured model size has 49.3949 MB, while the test-data memory requirement has approximately 82.1979 MB. The classification operation required 1.1118 s for the complete 261,978-record test set, while probability prediction required 1.0983 s. The combined inference time have 2.2101s. The resulting classification throughput have approximately 235,630 records/s, and the combined throughput have approximately 118,538 records/s. The average prediction time have approximately 0.004244 ms per record, presented in Table 14.

Table 13: Computational Performance of the Proposed RF Model

Computational Measure	Results
Test Records	261978
Selected Features	40
RF trees	200
Model size	49.3949 MB
Test-data memory	82.1979 MB
Classification prediction time	1.1118 s
Probability prediction time	1.0983 s
Total inference time	2.2101 s
Classification throughput	235629.95 records/s
Combined throughput	118537.72 records/s
Prediction time per record	0.004244 ms

These measurements demonstrate that the trained method have processed a large number of network-flow observations efficiently under the assessment computing environment.

The computational results are relevant to PA cybersecurity because agricultural IoT environments may require continuous monitoring of large volumes of network traffic. Prior research has also emphasized computational efficiency and lightweight security mechanisms for agricultural IoT and edge environments [5-6]. However, the current measurements have obtained under the experimental environment described in Section 4.4 and have not be interpreted as direct evidence of equivalent performance on resource-constrained agricultural gateways or embedded devices.

All 261,978 test observations have assigned risk categories and corresponding mitigation priorities, establishing a clear detection-to-risk-to-response pathway. The framework therefore extends conventional intrusion detection by integrating explainability, risk quantification, prioritization, and mitigation mapping. The computational results, including approximately 235,630 records/s classification throughput, further indicate efficient processing under the evaluated environment. However, the single-dataset evaluation, binary classification, deterministic mitigation mapping, and lack of real-world deployment validation indicate the need for further experimental validation.

6. Comparative Analysis

The experimental outcomes demonstrate that the proposed CSRM framework extends conventional agricultural IDSs by establishing a computational way from network-flow analysis to risk prioritization and mitigation response. Although the RF detector obtained high binary classification performance, the key contribution lies in integrating threat detection, probability estimation, SHAP-based explainability, quantitative risk scoring, risk-level classification, and deterministic mitigation mapping. This section compared the proposed framework with existing methods, highlights its benefits and limitations, and discusses its practical implications for PA cybersecurity.

Table 14: Comparative analysis of the proposed framework with existing approaches

Ref.	Technique / Approach	Dataset	AR (%)	PS (%)	RC (%)	FS (%)	ROC-AUC	Key Characteristic
[8]	Farm-Flow RF baseline	Farm-Flow	76.3700	81.7000	72.3000	72.3000	NR	Agriculture-specific network-flow IDS
[9]	CNN-DNN + IWHO	Farm-Flow	99.6700	NR	NR	92.1300	93.5200	Hybrid optimized deep learning
[10]	Lightweight DNN	18 IoT-IDS datasets	99.1400	NR	NR	NR	NR	Lightweight, dataset-agnostic IDS
[16]	SCAE-AGRU	Smart Agriculture IoT	99.9000	99.9000	99.7000	99.8000	NR	Privacy-preserving anomaly/IDS
[22]	FICNet (multiclass)	Farm-Flow, 8 classes	81.2500	91.2600	82.6200	80.4300	96.7800	Multiclass agricultural attack detection
Proposed Method	CSRM Framework	Farm-Flow	99.9282	99.9434	99.8811	99.9122	99.9519	Attack Detection + SHAP + Risk Assessment + Mitigation

In Table 15 compares the proposed CSRMF with representative recent cybersecurity and intrusion-detection approaches. The comparison shown that the proposed method achieved consistently high binary classification performance on the Farm-Flow dataset, with 99.8282% AR, 99.9434% PS, 99.8811% RC, 99.9122% FC, and 99.9519% ROC-AUC. The Farm-Flow RF baseline reported substantially lower values of 76.37% AC, 81.70% PS, 72.30% RC, and 72.30% FS [4]. The CNN-DNN + IWHO approach achieved 99.67% AC, 92.13% FS, and 93.52% ROC-AUC on Farm-Flow [5]. These results show that the proposed detector provides highly competitive performance on the same agriculture-specific benchmark, with improvements over the reported Farm-Flow RF baseline of 23.4582% points in accuracy, 18.2434% points in precision, 27.5811% points in recall, and 27.6122% points in F1-score. Compared with CNN-DNN + IWHO, the proposed method achieves a 0.1582%-point higher accuracy, although direct comparison have be interpreted cautiously because the reported studies may use different experimental configurations and evaluation protocols.

The comparison with FICNet further illustrates the importance of distinguishing binary and multiclass evaluation settings. FICNet reported 81.25% accuracy, 91.26% precision, 82.62% recall, 80.43% F1-score, and 96.78% ROC-AUC for its eight-class Farm-Flow classification setting [8,14]. These multiclass results are not directly comparable with the proposed binary Normal-versus-Attack formulation. Similarly, the Lightweight DNN study reported 99.14% average accuracy across 18 IoT-IDS datasets [6], which represents a broader multi-dataset evaluation rather than a Farm-Flow-specific result. The SCAE-AGRU approach

reported 99.90% AC, 99.90% PS, 99.70% RC, and 99.80% FS in a smart-agriculture IoT setting [9]. Therefore, the numerical comparison have interpreted in the context of the respective datasets and task formulations rather than as a universal ranking of algorithms.

After predictive performance, the main distinction of the proposed CSRM framework has integrated cybersecurity functionality. Existing methods in the comparison mainly emphasize intrusion or anomaly detection, whereas the proposed framework extends the detection process through attack-probability estimation, SHAP-based explainability, quantitative risk scoring, risk-level classification, and deterministic mitigation mapping. Thus, the contribution of the proposed CSRM framework has not solely its high classification performance but the establishment of a computational way from threat detection to interpretable risk prioritization and security response. This provides a more comprehensive cybersecurity decision-support capability for PA settings than a detection-only architecture.

The proposed CSRM framework integrated AI-based threat detection, attack-probability estimation, SHAP explain-ability, risk scoring, risk classification, and mitigation into a unified pipeline. The framework converts method predictions into a 0 – 100 risk score and classifies events as various levels for prioritized response. On FF dataset, the RF method achieved 235,630 records/s for classification and 118,538 records/s for combined classification and probability prediction. Despite, assessment has limited to binary classification on a single dataset, without practical deployment, various-class or zero-day validation. The 97.7185% Critical-risk proportion has FF dataset and formulation-dependent and have not been interpreted as practical threat prevalence. Similarly, SHAP provided predictive attribution rather than causal evidence. The framework therefore offers a structured decision-support way from detection to response, while deployment-specific evaluation of latency, memory, energy, and edge performance remains necessary.

7. Conclusion and Future Work

This study proposed a CSRM framework for PA to address the enhancing cybersecurity task associated with network-connected farming systems. The proposed framework integrates AI-based threat detection, attack-probability estimation, SHAP-based explainability, quantitative risk assessment, and risk-based mitigation within a unified workflow. The framework has implemented and evaluated through the FF agriculture dataset under a stratified binary classification setting. The RF-based detection method demonstrated strong predictive performance, obtained 99.9282% AC, 99.9434% PS, 99.8811% RC, 99.9122% FS, and 0.999519 ROC-AUC. The method also obtained an AP of 0.999982 and MCC of 0.958845, shown robust classification performance under the evaluated experimental conditions. The confusion-matrix analysis outcomes in a low overall error rate of 0.1718%, with a FNR of only 0.1189%. Beyond threat detection, the framework converts detection probabilities and severity information into a quantitative risk score, which have classified into various levels. These CSR levels have mapped to corresponding mitigation actions, including routine monitoring, enhanced monitoring, immediate investigation, and immediate incident response. In the test set, 97.7185% of records have classified as Critical risk data, demonstrating the ability of the proposed CSRM framework to prioritized detected security events according to their assessed risk. SHAP analysis further identifies influential network-flow characteristics, improving the interpretability of the AI-based detection process.

Future work will be extended the framework various class attack classification, cross agriculture -datasets to validation, and zero-day attack detection in PA and SA.

SDG ALIGNMENT

This research contributes to SDG 2 (Zero Hunger), SDG 9 (Industry, Innovation and Infrastructure), SDG 12 (Responsible Consumption and Production), SDG 13 (Climate Action), and SDG 15 (Life on Land) by promoting secure, sustainable, and AI- and IoT-enabled precision agriculture.

SDG Keywords: SDG 2, SDG 9, SDG 12, SDG 13, SDG 15, Precision Agriculture, IoT, AI, Cybersecurity, Sustainable Agriculture.

Acknowledgment

Author would like to acknowledge the MCN NO: IU/R&D/2026-MCN0005107

References

- [1]. Alahe, M. A., Wei, L., Chang, Y., Gummi, S. R., Kemesi, J., Yang, X., ... & Sher, M. (2024). Cyber security in smart agriculture: Threat types, current status, and future trends. *Computers and Electronics in Agriculture*, 226, 109401.
- [2]. Campoverde-Molina, M., & Luján-Mora, S. (2025). Cybersecurity in smart agriculture: A systematic literature review. *Computers & Security*, 150, 104284.
- [3]. Carvalho, N., Adão, T., Morais, R., Costa, A. R., & Peres, E. (2025). Cybersecurity in Precision Agriculture: a short review and a practical status assessment over mySense IoT-based platform. *Procedia Computer Science*, 256, 255-266.
- [4]. Kumar, S., & Faisal, M. (2024, November). Critical review of cybersecurity risks in precision agriculture. In *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)* (pp. 1-7). IEEE.
- [5]. Kumar, S., & Faisal, M. (2026). Analyzing and Mitigation of Cybersecurity Threats and Risks in Precision Agriculture Through IoT and AI Techniques. *International Journal of Computer Information Systems and Industrial Management Applications*, 18(17s), 95-113.
- [6]. Ansar, S. A., Kumar, S., Khan, M. W., Yadav, A., & Khan, R. A. (2020). Enhancement of two-tier ATM security mechanism: towards providing a real-time solution for network issues. *International Journal of Advanced Computer Science and Applications*, 11(7).
- [7]. Kumar, S., Srivastava, A., Chaurasiya, P. K., Kushawaha, A., & Vishal, V. (2022, December). DCT and SVD-based watermarking technique for imperceptibility and robustness of medical images. In *2022 4th international conference on advances in computing, communication control and networking (ICAC3N)* (pp. 2335-2339). IEEE.
- [8]. Ferreira, R., Bispo, I., Rabadão, C., Santos, L., & Costa, R. L. D. C. (2025). Farm-flow dataset: Intrusion detection in smart agriculture based on network flows. *Computers and Electrical Engineering*, 121, 109892.
- [9]. Muhammed, D., Ahvar, E., Ahvar, S., Trocan, M., Montpetit, M. J., & Ehsani, R. (2024). Artificial Intelligence of Things (AIoT) for smart agriculture: A review of architectures, technologies and solutions. *Journal of Network and Computer Applications*, 228, 103905.
- [10]. Villafranca, A., & Cano, M. D. (2025). A lightweight edge-DL intrusion detection system for IoT sustainable smart-agriculture. *Internet of Things*, 101818.
- [11]. Dwivedi, S., Afaq, S. A., Faisal, M., Siddique, M., & Verma, N. K. (2025). Transforming Healthcare: Exploring the Intersection of IoT and Machine Learning. In *Startup-Driven E-Government: Digital Innovation for Sustainable Ecosystems* (pp. 561-602). IGI Global Scientific Publishing.
- [12]. Kumar, S., Singh, S., & Chaurasia, P. K. (2023, November). Enhancing security of medical image transmission: An innovative fuzzy-AHP approach. In *International Conference on Trends in Computational and Cognitive Engineering* (pp. 471-483). Singapore: Springer Nature Singapore.
- [13]. Kumar, S., Kumar, A., Kumar, S., & Chaurasia, P. K. (2023, September). Comprehensive analysis of cloud security: Issues & challenges. In *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)* (Vol. 6, pp. 622-627). IEEE.
- [14]. Ferreira, R., Bispo, I., Rabadão, C., Santos, L., & Costa, R. L. D. C. (2025). Farm-flow dataset: Intrusion detection in smart agriculture based on network flows. *Computers and Electrical Engineering*, 121, 109892.
- [15]. Juboraj, M. F. U. A., Al Farid, F., Zabin, M., Uddin, J., Hossain, M. I., & Mansor, S. (2026). FICNet: A Deep Learning Framework for Intrusion Detection in Agricultural Internet of Things. *Computer Modeling in Engineering & Sciences*, 148(1), 48.
- [16]. Ferrag, M. A., Shu, L., Friha, O., & Yang, X. (2021). Cyber security intrusion detection for agriculture 4.0: Machine learning-based solutions, datasets, and future directions. *IEEE/CAA Journal of Automatica Sinica*, 9(3), 407-436.
- [17]. Shrivastava, A., Praveen, R. V. S., Vemuri, H. K., Peri, S. S. S. R. G., Sista, S., & Hasan, M. M. (2027). Future Directions and Challenges in Smart Agriculture and Cybersecurity. *Sustainable Agriculture Production Using Blockchain Technology*, 265-276.
- [18]. Manono, B. O., Mwami, B., Mutavi, S., & Nzilu, F. (2026). Precision farming with smart sensors: Current state, challenges and future outlook. *Sensors*, 26(3), 882.
- [19]. Ahmad, M., Pratap, R., Maurya, S. K., Iqbal, M. S. A., & Kumar, S. (2026). A Unified Framework for Classifying Big Data Security Algorithms: A Healthcare Perspective. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, 96(1), 127-137.

- [20]. Muganthan, S., Srivignesh, L., Rajashree, S., Miriam, A. J., & Muneera, M. N. (2026, March). Real-Time Cyber Threat Intelligence Detection and Response Using Deep Learning for Advanced Threat Classification. In 2026 3rd International Conference on Emerging Trends in Engineering and Medical Sciences (ICETEMS) (pp. 1-7). IEEE.
- [21]. Villafranca, A., & Cano, M. D. (2025). A lightweight edge-DL intrusion detection system for IoT sustainable smart-agriculture. *Internet of Things*, 101818.
- [22]. Juboraj, M. F. U. A., Al Farid, F., Zabin, M., Uddin, J., Hossain, M. I., & Mansor, S. (2026). FICNet: A Deep Learning Framework for Intrusion Detection in Agricultural Internet of Things. *Computer Modeling in Engineering & Sciences*, 148(1), 48.
- [23]. Xie, M., Liu, J., Chen, S., Lin, M., Xu, G., Xu, Z., & Chang, Z. (2025). Traceability and identity protection in smart agricultural IoT system framework based on blockchains. *IEEE Transactions on Dependable and Secure Computing*, 22(5), 5335-5351.
- [24]. Kumar, S., Ahmad, M., Maurya, S. K., Pratap, R., Chaurasia, P. K., & Khan, R. A. (2025). Soft computing-driven framework for enhanced security in medical image transmission. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, 95(2), 177-189.
- [25]. Mahlous, A. R. (2024). Security analysis in smart agriculture: Insights from a cyber-physical system application. *Computers, Materials & Continua*, 79(3), 4781-4803.
- [26]. Afaq, S. A., Uzma, S., & Yasmeen, G. (2023). The Critical Impact of Cyber Threats on Digital Economy. In *Advances in Cyberology and the Advent of the Next-Gen Information Revolution* (pp. 86-108). IGI Global Scientific Publishing
- [27]. Ahamad, M.K. and Bharti, A.K., 2021, March. Prevention from COVID-19 in India: fuzzy logic approach. In 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 421-426). IEEE.
- [28]. Kumar, S., Devi, M., Singh, S., Chaurasia, P. K., & Khan, R. A. (2023, September). Prioritization of medical image security features: fuzzy AHP approaches. In 2023 6th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 6, pp. 540-545). IEEE.
- [29]. Rudrakar, S., Rughani, P., & Sadineni, L. (2025). Digital forensics and incident response management model for IoT based agriculture. *Scientific Reports*, 15(1), 17797.
- [30]. Keerthna, G., Jayamala, C., & Lavaraju, B. (2027). Risk Assessment and Cybersecurity Strategies for Agricultural Systems. *Sustainable Agriculture Production Using Blockchain Technology*, 203-214.
- [31]. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.
- [32]. Juboraj, M. F. U. A., Al Farid, F., Zabin, M., Uddin, J., Hossain, M. I., & Mansor, S. (2026). FICNet: A Deep Learning Framework for Intrusion Detection in Agricultural Internet of Things. *Computer Modeling in Engineering & Sciences*, 148(1), 48.