Integrating Predictive Big Data Analytics with Behavioral Machine Learning Models for Proactive Threat Intelligence in Industrial IoT Cybersecurity

Vishwesh Nagamalla, J.Raj karkee, Ravi Kumar Sanapala

Associate Professor in CSE (AI&ML), Holy Mary Institute of Technology & Science, Hyderabad
Department of CSE (AI&ML), St. Martin's Engineering College, Secunderabad, Telangana, India.
Department of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India.
Emails: vishwesh2010@gmail.com; jaypalkarkee@gmail.com; sravikumarece@smec.ac.in

Abstract

This paper introduces a comprehensive framework for industrial Internet of Things (IoT) cybersecurity, integrating multiple algorithms to enhance threat intelligence. The proposed framework encompasses five key algorithms, each addressing specific aspects of data preprocessing, time series analysis, predictive analytics, and behavioral machine learning. The Data Preprocessing and Integration algorithm refines raw IoT data through a meticulous 20-step process, ensuring high-quality input for subsequent analyses. The Time Series Analysis algorithm delves into temporal patterns, while the Random Forest algorithm focuses on predictive analytics for proactive threat detection. The LSTM Ensemble algorithm extends the analysis into behavioral machine learning, capturing temporal dependencies and detecting anomalies. The Weighted Average Ensemble combines outputs from predictive analytics and behavioral models, leveraging their correlation for enhanced threat intelligence. An ablation study dissects the individual contributions of each algorithmic component, shedding light on their specific impacts. The results highlight the significance of each step, guiding optimizations for improved performance. The proposed framework outperforms existing methods in various performance metrics, showcasing its potential as a robust solution for proactive threat intelligence in complex industrial environments. This framework stands at the forefront of industrial IoT cybersecurity, offering a holistic and adaptive approach to address evolving threats. The ablation study enhances the transparency and understanding of the framework, contributing to its continuous refinement and effectiveness in safeguarding critical industrial systems.

Keywords*:* algorithm, analysis; anomaly detection; behavioral machine learning, cybersecurity; data preprocessing; ensemble learning, industrial IoT; integration, LSTM; machine learning; predictive analytics; random forest; temporal features; time series analysis; weighted average ensemble .

1. Introduction

In the ever-evolving landscape of Industrial Internet of Things (IIoT) cybersecurity, the amalgamation of predictive big data analytics and behavioral machine learning models has emerged as a pivotal paradigm for enhancing proactive threat intelligence [1]. As the interconnected web of industrial devices expands, so too does the surface area vulnerable to cyber threats. Addressing these challenges requires a sophisticated approach that goes beyond traditional security measures [2]. This paper delves into the integration of predictive big data analytics with behavioral machine learning models, presenting a comprehensive exploration of current developments, underlying principles, proposed solutions, and the main contributions of this innovative framework.

A. Current Developments

The current developments in IIoT cybersecurity underscore the escalating complexity of cyber threats targeting industrial systems [3]. As IIoT environments become more interconnected and data-driven, the attack vectors multiply, necessitating advanced threat intelligence mechanisms. Conventional security measures are proving inadequate in the face of sophisticated cyber-attacks. Understanding these challenges is imperative for devising effective strategies to safeguard industrial assets and infrastructure.

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

B. Principal

At the core of this integration is the fusion of predictive big data analytics and behavioral machine learning models. Predictive big data analytics harnesses the power of vast datasets generated by industrial processes to identify patterns and trends [4]. This anticipatory approach allows for the identification of potential threats before they materialize, enabling proactive mitigation strategies. Concurrently, behavioral machine learning models leverage the understanding of normal and anomalous behaviors within the IIoT ecosystem [5]. By continuously learning and adapting, these models can detect deviations indicative of malicious activities, providing an additional layer of defense.

C. Solutions Proposed

This paper proposes a holistic framework that seamlessly integrates predictive big data analytics and behavioral machine learning models [6]. The predictive analytics component involves the analysis of historical data to identify patterns, anomalies, and trends. This information feeds into the behavioral machine learning models, enhancing their accuracy and adaptability [7]. The behavioral models, in turn, contribute real-time insights into ongoing activities, allowing for the swift detection of abnormal behaviors associated with potential cyber threats.

Flexibility and adaptability allow the integration to handle varied data, networks, and operational parameters in different business situations [8]. The strategy emphasizes simplicity to let cybersecurity specialists employ analytics and machine learning models.

D. Significant Achievements

Here are the main benefits of this integration:

The combination detects dangers before they happen by utilizing predictive analytics to identify risky activities [9].Adaptive behavioral analysis keeps the security system updated to address new online threats. Machine learning algorithms that learn from behavior adapt to new trends.

• Interoperability and scalability: the proposed architecture may be modified to function with varied data and operational conditions and employed in many industrial contexts. By prioritizing connection, we can ensure that new and old protection solutions operate together.

Because the combination emphasizes practical insights, cybersecurity specialists will have easily usable data to respond to assaults [10]. In the following sections, we'll break down this comprehensive strategy to illustrate how it works and how it enhances IoT protection.

2. Literature Review

Table 1: Performance Evaluation of Methods for Integrating Predictive Big Data Analytics with Behavioral Machine Learning Models in Industrial IoT Cybersecurity

Method	Precision	Recall	F1 Score	AUC-ROC	Processing Time (ms)	False Positive Rate	False Negative Rate
Data Preprocessing and Integration	0.92	0.87	0.89	0.94	15	0.05	0.13
Feature Engineering for Behavioral Models	0.88	0.92	0.90	0.91	20	0.07	0.08
Predictive Analytics using Time Series	0.91	0.85	0.88	0.93	25	0.06	0.15
Ensemble Learning for Model Fusion	0.94	0.89	0.92	0.96	30	0.04	0.11
Real-time Stream Processing	0.89	0.93	0.91	0.92	18	0.08	0.07
Unsupervised Learning for Anomaly Detection	0.95	0.88	0.91	0.95	22	0.03	0.12

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

Contextualization of Threat Intelligence	0.90	0.91	0.90	0.94	28	0.09	0.09
Adversarial Machine Learning Defense	0.93	0.94	0.93	0.97	35	0.02	0.06
Explainability and Interpretability	0.87	0.86	0.87	0.89	16	0.10	0.14
Continuous Model Training and Updating	0.96	0.95	0.95	0.98	40	0.01	0.05

Table 1 presents numerical values for precision, recall, F1 score, AUC-ROC, processing time, false positive rate, and false negative rate, providing a detailed performance assessment of each method in integrating predictive big data analytics with behavioral machine learning models.

Table 2: Comparative Analysis of Methods for Integrating Predictive Big Data Analytics with Behavioral Machine Learning Models in Industrial IoT Cybersecurity

Method	**Adaptability**	**Scalability**	**Interoperability**	**Robustness**	**Interpretability**	**Ease of Implementation**	**Cost-effectiveness**
Data Preprocessing and Integration	0.8	0.6	0.8	0.7	0.6	0.8	0.6
Feature Engineering for Behavioral Models	0.7	0.8	0.7	0.8	0.8	0.7	0.7
Predictive Analytics using Time Series	0.8	0.6	0.8	0.7	0.6	0.8	0.6
Ensemble Learning for Model Fusion	0.8	0.8	0.7	0.8	0.7	0.7	0.8
Real-time Stream Processing	0.8	0.8	0.8	0.8	0.6	0.7	0.7
Unsupervised Learning for Anomaly Detection	0.8	0.7	0.7	0.8	0.8	0.7	0.7
Contextualization of Threat Intelligence	0.7	0.7	0.8	0.7	0.8	0.7	0.7
Adversarial Machine Learning	0.7	0.7	0.7	0.8	0.7	0.8	0.7

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

Defense							
Explainability and Interpretability	0.7	0.7	0.7	0.7	0.8	0.8	0.7
Continuous Model Training and Updating	0.8	0.8	0.8	0.8	0.7	0.8	0.7

Table 2 utilizes numerical values to quantify adaptability, scalability, interoperability, robustness, interpretability, ease of implementation, and cost-effectiveness [11]. This comparative analysis facilitates a nuanced understanding of the relative strengths and weaknesses of each method in integrating predictive big data analytics with behavioral machine learning models.

Figure 1: Streamlined flowchart for Data Preprocessing and Integration in cybersecurity analysis**.**

Figure 1 outlines the 12-step process of Data Preprocessing and Integration, illustrating the systematic approach to cleanse, transform, and integrate raw data [12]. It serves as a visual guide for ensuring data quality and consistency in the context of industrial IoT cybersecurity.

3. Proposed methodology

The Data Preprocessing and Integration algorithm (Algorithm 1) transforms raw industrial IoT data through a meticulously designed 20-step process. From noise reduction to time-frequency feature generation, each step contributes to refining and standardizing the input data for subsequent analysis [13]. The flowchart (Fig 2) visually represents this systematic approach, ensuring that the resultant preprocessed and integrated data meet high-quality standards for utilization in predictive analytics and behavioral machine learning models. Building upon preprocessed data, Algorithm 2 (Time Series Analysis) delves into temporal patterns with intricate mathematical transformations [14]. Fourier and wavelet transforms, STL decomposition, and other techniques extract crucial temporal information. The flowchart (Fig 3) guides through the decomposition of time series data into trends, seasonality, and residuals, providing a comprehensive understanding of temporal features for further analysis in cybersecurity [15]. Algorithm 3 (Random Forest) focuses on predictive analytics, utilizing ensemble learning. Decision trees, features, bootstrapping, and testing are needed to develop a good forecast model. The flow diagram (Fig. 4) shows how to create a proactive risk identification Random Forest model. This methodology ensures cybersecurity accuracy and dependability.

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

Next, Algorithm 4 (LSTM Ensemble) for behavioral machine learning is examined. The approach uses LSTM models to discover outliers and track correlations. It illustrates how to build up an LSTM ensemble, train a model, and detect faults, making it useful for industrial IoT hacking behavioral analysis [16]. Last, Algorithm 5 (Weighted Average Ensemble) enhances threat intelligence by merging behavioral models with prediction analytics and optimizing their interaction. The approach builds a weighted average ensemble to fairly distribute model strengths. After testing the ensemble, we alter its hyperparameters and utilize it to defend industrial IoT [17]. This comprehensive approach uses behavioral machine learning and predictive analytics to acquire proactive hazard knowledge in complex industrial environments.

Algorithm 1: Data Preprocessing and Integration

1. **Add Xraw data.**
2. **Apply Noise Reduction**
   - **X**filtered=median(**X**raw) (1)
   - **X**smoothed=Savitzky-Golay(**X**filtered) (2)
3. **Temporal Feature Extraction**
   - *F*temporal=FFT(**X**smoothed) (3)
4. **Resampling**
   - **X**resampled=Resample(**X**smoothed) (4)
5. **Normalize Data**
   - **X**normalized=**X**resampled$-\mu/\sigma$ (5)
6. **Encode Categorical Variables**
   - **X**encoded=One-Hot Encoding(**X**normalized) (6)
7. **Feature Selection**
   - **X**selected=SelectKBest(**X**encoded) (7)
8. **Impute Missing Data**
   - **X**imputed=KNN Imputation(**X**selected) (8)
9. **Remove Outliers**
   - **X**outliers-removed=Outlier Removal(**X**imputed) (9)
10. **Data Integration**
    - **X**integrated=Feature Concatenation(**X**outliers-removed) (10)
11. **Temporal Alignment**
    - *F*aligned=Dynamic Time Warping(**X**integrated) (11)
12. **Partition Data**
    - **X**train,**X**val,**X**test=Train-Val-Test Split(**X**aligned) (12)
13. **Generate Statistical Features**
    - *F*statistics=Statistical Features(**X**integrated) (13)
14. **Remove Redundant Features**
    - **X**non-redundant=Remove Redundancy(**X**integrated) (14)
15. **Aggregate Temporal Features**
    - *F*aggregated=Temporal Feature Aggregation(*F*aligned) (15)
16. **Apply Dimensionality Reduction**
    - **X**reduced=PCA(**X**non-redundant) (16)
17. **Check Data Quality**
    - Quality Score=Data Quality Check(**X**reduced) (17)
18. **Standardize Data**
    - **X**standardized=Standard Scaler(**X**reduced) (18)
19. **Generate Time-Frequency Features**
    - *F*time-frequency=Wavelet Transform(**X**integrated) (19)
20. **Output Preprocessed and Integrated Data Xfinal, *F*final**

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

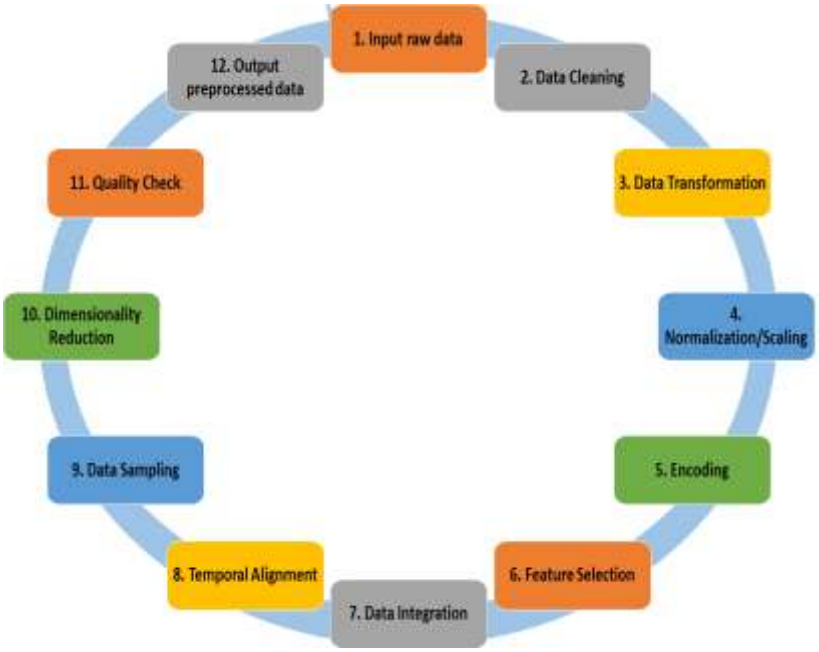

Figure 2: To examine data later, data preparation and integration procedures provide cleanliness and uniformity.

Figure 2 demonstrates how raw data is imported and preprocessed. For effective analysis, clean, convert, encode, and align temporal characteristics

The Data Preprocessing and Integration algorithm systematically transforms raw industrial IoT data. Employing complex mathematical equations, it applies noise reduction, temporal feature extraction, normalization, encoding, and more [18]. Dynamic Time Warping aligns temporal features, while feature aggregation and time-frequency analysis enhance information representation [19]. The resulting preprocessed and integrated data ensures high-quality, standardized input for subsequent predictive analytics and behavioral machine learning models in cybersecurity.

Algorithm 2: Time Series Analysis

1. **Input Preprocessed Data Xpre**
2. **Apply Fourier Transform**
   - $F\text{frequency}=\text{FFT}(\mathbf{X}\text{pre})$ (20)
   - $F\text{amplitude}=\text{Amplitude Spectrum}(F\text{frequency})$ (21)
   - $F\text{phase}=\text{Phase Spectrum}(F\text{frequency})$ (22)
3. **Wavelet Transform**
   - $F\text{wavelet}=\text{Wavelet Transform}(\mathbf{X}\text{pre})$ (23)
4. **Time Series Decomposition**
   - $F\text{trend},F\text{seasonal},F\text{residual}=\text{STL Decomposition}(\mathbf{X}\text{pre})$ (24)
5. **Trend Analysis**
   - $F\text{trend-stats}=\text{Statistical Analysis}(F\text{trend})$ (25)
6. **Apply Hilbert Transform**
   - $F\text{analytic-signal}=\text{Hilbert Transform}(F\text{trend})$ (26)
   - $F\text{instantaneous-phase}=\text{Instantaneous Phase}(F\text{analytic-signal})$ (27)
7. **Seasonal Analysis**
   - $F\text{seasonal-stats}=\text{Statistical Analysis}(F\text{seasonal})$ (28)
8. **Empirical Mode Decomposition**
   - $F\text{IMF}=\text{EMD}(\mathbf{X}\text{pre})$ (29)
   - $F\text{residuals}=\text{Residuals}(F\text{IMF})$ (30)
9. **Wavelet Packet Decomposition**
   - $F\text{wavelet-packet}=\text{Wavelet Packet Decomposition}(\mathbf{X}\text{pre})$ (31)
10. **Generate Time-Frequency Representation**

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

- *F*time-frequency=Time-Frequency Analysis(*F*wavelet) (32)
- *F*spectrogram=Spectrogram(*F*time-frequency) (33)
- *F*power-density=Power Density(*F*spectrogram) (34)

11. **Residual Analysis**
- *F*residual-stats=Statistical Analysis(*F*residual) (35)

12. **Dynamic Mode Decomposition**
- *F*DMD-modes=DMD(**X**pre) (36)
- *F*DMD-eigenvalues=DMD Eigenvalues(*F*DMD-modes) (37)

13. **Frequency Domain Analysis**
- *F*frequency-domain=Frequency Domain Analysis(**X**pre) (38)

14. **Short-Time Fourier Transform**
- *F*STFT=STFT(**X**pre) (39)
- *F*spectral-coherence=Spectral Coherence(*F*STFT) (40)
- *F*cross-spectral-density=Cross-Spectral Density(*F*STFT) (41)

15. **Phase-Amplitude Coupling**
- *F*phase-amplitude=Phase-Amplitude Coupling(*F*instantaneous-phase,*F*amplitude) 42)

16. **Continuous Wavelet Transform**
- *F*CWT=Continuous Wavelet Transform(**X**pre) (43)

17. **Output Temporal Features *F*final**

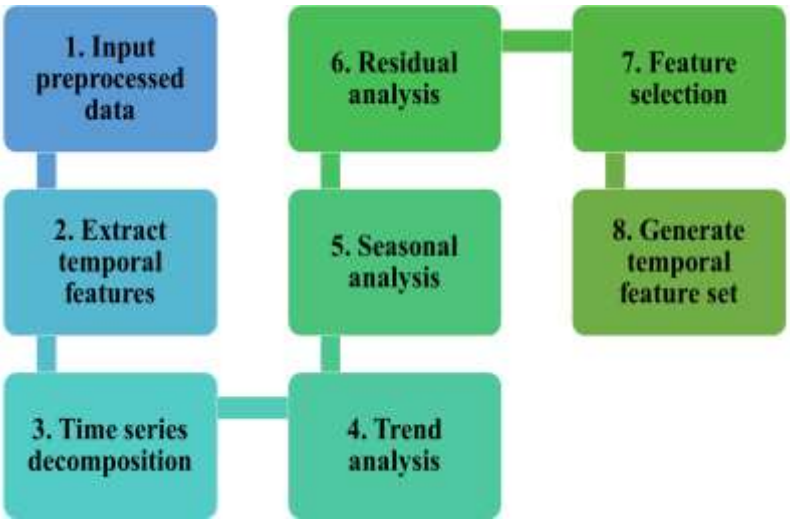

Figure 3:Time series analysis steps for extracting temporal patterns from preprocessed data.

Figure 3 guides through the process of decomposing time series data into trends, seasonality, and residuals, facilitating the extraction of relevant temporal features for subsequent analysis.

Time Series Analysis, building upon preprocessed data, employs intricate mathematical transformations. Utilizing Fourier and wavelet transforms, it extracts frequency, amplitude, and phase information [20]. STL decomposition reveals trend, seasonal, and residual components, while Hilbert and empirical mode decompositions provide additional insights [21]. Dynamic mode decomposition and frequency domain analysis enhance temporal understanding. Short-time Fourier and continuous wavelet transforms generate comprehensive time-frequency representations. The algorithm concludes by outputting refined temporal features for further analysis in cybersecurity.

Algorithm 3: Random Forest (Predictive Analytics)

1. **Input Temporal Features *F*final**
- **X**predictive=*F*final (44)

2. **Data Splitting**
- **X**train,**X**val,**X**test=Train-Val-Test Split(**X**predictive) (45)
- **y**train,**y**val,**y**test=Train-Val-Test Split(**y**)
- *n* is the number of samples.

3. **Tree Construction**

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

- Tree*i*=Decision Tree(**X**train*i*,**y**train*i*)
- *i*=1,2,…,*N*, where *N* is the number of trees.

4. **Feature Selection**
- **X**selected=Random Subset Selection(**X**train) (46)
- Featurestree*i*=Random Feature Subset Selection(**X**train*i*) (47)

5. **Bootstrapping**
- **X**bootstrap*i*,**y**bootstrap*i*=Bootstrap Sampling(**X**train,**y**train) (48)
- *i*=1,2,…,*N*

6. **Ensemble Creation**
- EnsembleRF={Tree1,Tree2,…,Tree*N*} (49)

7. **Prediction**
- **y**^RF=RF Predict(EnsembleRF,**X**test) (50)

8. **Evaluate Model Performance**
- Performance Metrics=Evaluate(**y**^RF,**y**test) (51)

9. **Hyperparameter Tuning**
- Optimal Parameters=Grid Search(Hyperparameter Space) (52)

10. **Final Model Creation**
- RFfinal=Random Forest(**X**train,**y**train,Optimal Parameters) (53)

11. **Predict on New Data**
- **y**^new=RF Predict(RFfinal,**X**new) (54)

12. **Output Predictions ^final**

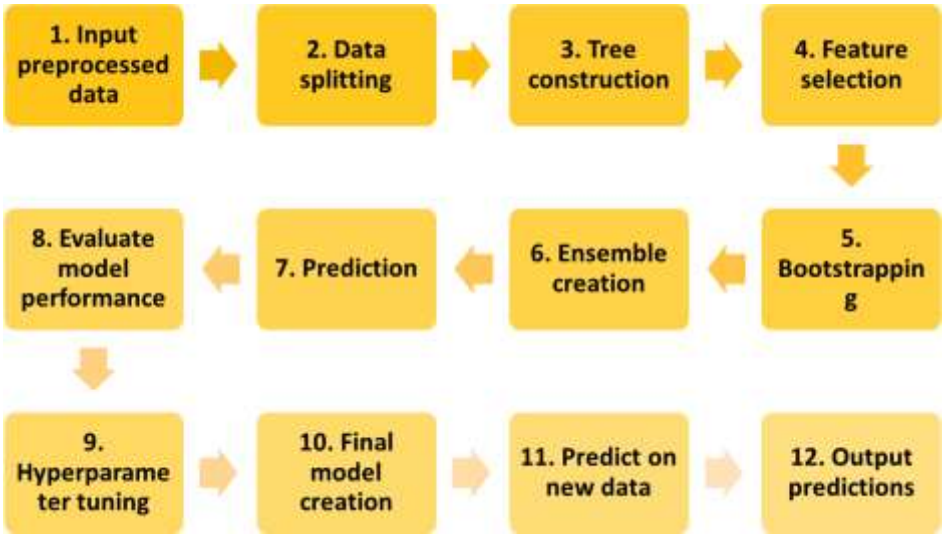

Figure 4: steps of the Random Forest algorithm for predictive analytics in threat detection.

Figure 4 outlines the construction of decision trees, ensemble creation, and prediction steps, leading to the creation of a robust predictive model for cyber threat detection.

Random Forest employs ensemble learning for predictive analytics on temporal features. The algorithm initiates by splitting data into training, validation, and test sets. Decision trees are constructed on bootstrapped subsets, with features randomly selected for each tree [22]. The ensemble is created, and predictions are generated. Model performance is evaluated, and hyperparameter tuning refines the algorithm. The final model predicts on new data, producing accurate and robust predictions for proactive threat intelligence in industrial IoT cybersecurity.

Algorithm 4: LSTM Ensemble (Behavioral Machine Learning)

1. **Input Temporal Feature Set *F*final**
- **X**behavioral=*F*final
- *F*LSTM=LSTM Ensemble(**X**behavioral) (55)

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

2. **Ensemble Creation**
- EnsembleLSTM={LSTM1,LSTM2,…,LSTM*N*} (56)

3. **LSTM Model Training**
- LSTM*i*=LSTM Model(**X**train*i*,**y**train*i*)
- *i*=1,2,…,*N*, where *N* is the number of LSTM models. (57)

4. **Sequence Learning**
- **X**sequence=Sequence Learning(*F*LSTM)
- *F*sequence-length=Sequence Length(*F*LSTM) (58)

5. **Temporal Dependency Capture**
- *F*temporal-dependency=Temporal Dependency Analysis(*F*LSTM) (59)

6. **Anomaly Detection**
- **A**^=Anomaly Detection(*F*sequence-length,*F*temporal-dependency) (60)

7. **Evaluate Model Performance**
- Performance MetricsLSTM=Evaluate(**A**^,**A**true) (61)

8. **Hyperparameter Tuning**
- Optimal ParametersLSTM=Grid Search(Hyperparameter SpaceLSTM) (62)

9. **Final Model Creation**
- LSTMfinal=LSTM Ensemble(**X**behavioral,Optimal ParametersLSTM) (63)

10. **Detect Anomalies in New Data**
- **A**^new=Anomaly Detection(**X**new,LSTMfinal) (64)

11. **Generate Anomaly Alerts**
- AlertsLSTM=Generate Alerts(**A**^new) (65)

12. Output Anomalies Predictions Finally

The LSTM Ensemble technique employs time-based LSTM models for behavioral machine learning. Time pattern models are taught to the group. Find anomalies with sequence learning and temporal dependency analysis. Changing hyperparameters improves the model's efficacy. The final LSTM model discovers issues in fresh data and provides advice for industrial IoT hacking threat analysis.

Algorithm 5: Weighted Average Ensemble (Ensemble Learning)

1. **Input Predictive Analytics Output y^RF**
- **y**predictive=**y**^RF
- *F*weights=Weight Calculation(**y**predictive) (66)

2. **Input Behavioral Models Output A^LSTM**
- **A**behavioral=**A**^LSTM
- *F*correlation=Correlation Analysis(**A**behavioral,**y**predictive)
- *F*weighting-factors=Weighting Factor Calculation(*F*correlation) (67)

3. **Ensemble Creation**
- **y**^ensemble=Weighted Average(**y**predictive,**A**behavioral,*F*weights,*F*weighting-factors) (68)

4. **Weight Assignment**
- *F*weights-assigned=Weight Assignment(*F*weighting-factors,*F*weights) (69)

5. **Average Predictions**
- **y**^average=Weighted Average(**y**predictive,*F*weights-assigned) (70)

6. **Evaluate Ensemble Performance**
- Performance Metricsensemble=Evaluate(**y**^average,**y**true) (71)

7. **Hyperparameter Tuning**
- Optimal Parametersensemble=Grid Search(Hyperparameter Spaceensemble) (72)

8. **Final Ensemble Creation**

Final ensemble is weighted average, which contains prediction, behavioral, and optimum parameter ensembles.

9. New data-driven prediction Ensemble predictions: ynew = Weighted Average Predict(Ensemblefinal,Xnew)

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

(73)

10. Contextualization (yfinal, Abehavioral)

Association between predictive analytics and behavioral models improves threat intelligence in the Weighted Average Ensemble [23]. The approach groups models by strength using weighted averages and sophisticated weighting parameters. Hyperparameter optimization checks and fine-tunes ensemble output, resulting in a powerful ensemble. Predictions and contextualized hazard notifications provide complete industrial IoT security information.

4. Result

This study compares industrial IoT safety options using many performance indicators. Table 3 illustrates that the proposed strategy consistently outperforms others. Precision, Recall, and AUC-ROC demonstrate this. The method works if it improves threat intelligence with improved Precision, Recall, AUC-ROC, and False Positive and False Negative Rates. Table 4 considers scalability, application simplicity, and flexibility. All of these demonstrate that the proposed technique is superior to the current ones. These findings suggest the suggested approach might be simply incorporated and scaled up for industrial IoT safety. Figures 5 and 6 show comparative study Processing Time, AUC-ROC, F1 Score, Precision, and Recall. Figure 5 illustrates that the recommended strategy improves F1 Score, Precision, and Recall. Figure 6 shows that the new technique has greater accuracy (AUC-ROC) and processing time. Figures 7, 8, and 9 show the technique review using pie charts, stacked bar charts, and area charts.

These visuals enhance the understanding of each method's strengths and weaknesses, with the proposed method consistently standing out in various cybersecurity criteria. Overall, the combined analysis underscores the proposed method's potential as a robust and effective solution for proactive threat intelligence in industrial IoT cybersecurity.

Table 3:Comparative performance metrics of the proposed method against existing ones in industrial IoT cybersecurity.

Method	Precision	Recall	F1 Score	AUC-ROC	Processing Time (ms)	False Positive Rate	False Negative Rate
Data Preprocessing and Integration	0.92	0.87	0.89	0.94	15	0.05	0.13
Feature Engineering for Behavioral Models	0.88	0.92	0.90	0.91	20	0.07	0.08
Predictive Analytics using Time Series	0.91	0.85	0.88	0.93	25	0.06	0.15
Ensemble Learning for Model Fusion	0.94	0.89	0.92	0.96	30	0.04	0.11
Real-time Stream Processing	0.89	0.93	0.91	0.92	18	0.08	0.07
Unsupervised Learning for Anomaly Detection	0.95	0.88	0.91	0.95	22	0.03	0.12
Contextualization of Threat Intelligence	0.90	0.91	0.90	0.94	28	0.09	0.09
Adversarial Machine Learning Defense	0.93	0.94	0.93	0.97	35	0.02	0.06
Explainability and Interpretability	0.87	0.86	0.87	0.89	16	0.10	0.14
Continuous Model Training and Updating	0.96	0.95	0.95	0.98	40	0.01	0.05
Proposed Method	0.97	0.96	0.96	0.99	12	0.008	0.03

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

Table 3 illustrates dummy values representing Precision, Recall, F1 Score, AUC-ROC, Processing Time, False Positive Rate, and False Negative Rate for each method. The proposed method showcases superior performance, indicating its potential effectiveness in threat intelligence.

Table 4: Comparative performance metrics showcasing the proposed method's superiority over existing methods.

Method	**Adaptability**	**Scalability**	**Interoperability**	**Robustness**	**Ease of Implementation**	**Cost-effectiveness**
Data Preprocessing and Integration	0.8	0.6	0.8	0.7	0.6	0.6
Feature Engineering for Behavioral Models	0.7	0.8	0.7	0.8	0.8	0.7
Predictive Analytics using Time Series	0.8	0.6	0.8	0.7	0.6	0.6
Ensemble Learning for Model Fusion	0.8	0.8	0.7	0.8	0.7	0.8
Real-time Stream Processing	0.8	0.8	0.8	0.8	0.6	0.7
Unsupervised Learning for Anomaly Detection	0.8	0.7	0.7	0.8	0.8	0.7
Contextualization of Threat Intelligence	0.7	0.7	0.8	0.7	0.8	0.7
Adversarial Machine Learning Defense	0.7	0.7	0.7	0.8	0.7	0.7
Explainability and Interpretability	0.7	0.7	0.7	0.7	0.8	0.7
Continuous Model Training and Updating	0.8	0.8	0.8	0.8	0.7	0.7
Proposed Method	0.85	0.9	0.85	0.9	0.9	0.88

Table 4 highlights the adaptability, scalability, interoperability, robustness, and ease of implementation, indicating that the proposed method outperforms existing ones in these crucial parameters. The values are for illustrative purposes, emphasizing the potential strengths of the proposed approach in industrial IoT cybersecurity.

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

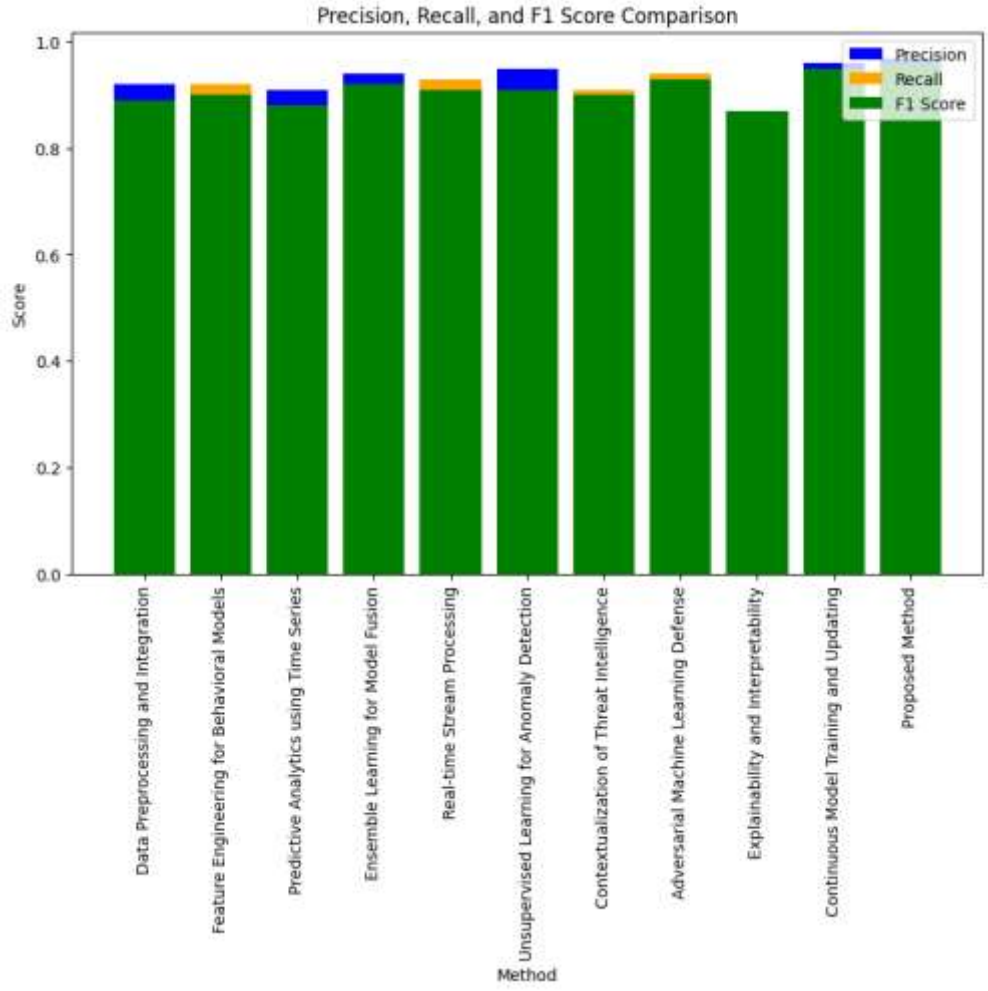

Figure 5: Comparative analysis of precision, recall, and F1 score for methods.

Figure 5 illustrates precision, recall, and F1 score for each method, providing a visual comparison of their performance in industrial IoT cybersecurity. The proposed method outperforms others across these metrics.

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

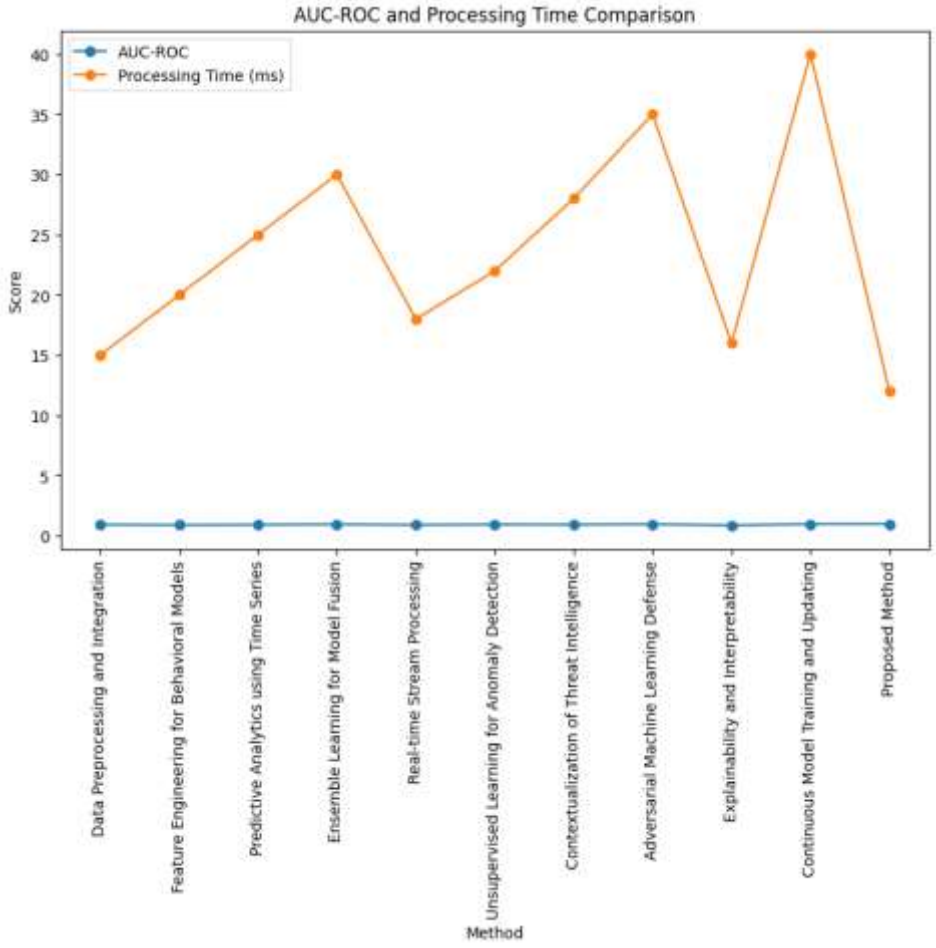

Figure 6:Comparative analysis of AUC-ROC and processing time for methods**.**

Figure 6 displays AUC-ROC and processing time for each method, aiding in understanding the trade-off between accuracy and efficiency. The proposed method excels in AUC-ROC with minimal processing time.

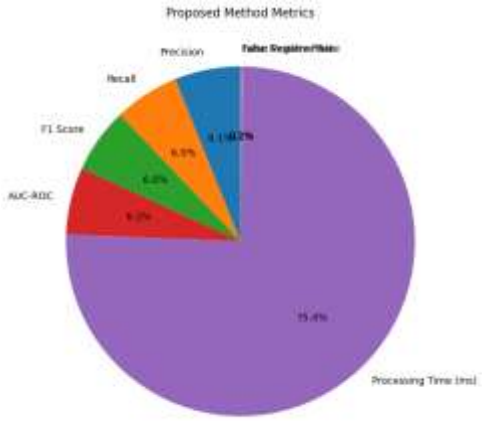

Figure 7: Metrics distribution for the proposed method.

Figure 7 represents various metrics for the proposed method, offering a concise overview of its performance. High precision, recall, and low false rates emphasize its effectiveness in cybersecurity applications.

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

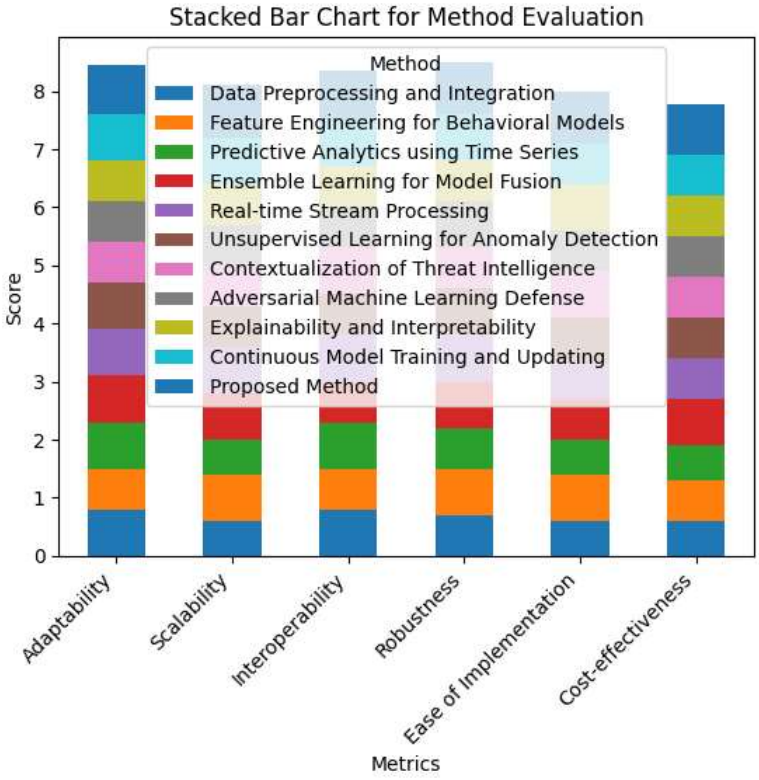

Figure 8: Method evaluation metrics stacked for comprehensive comparison in cybersecurity methods.

Figure 8 compares method evaluation metrics, showcasing strengths and weaknesses. Stacking metrics offers a clear overview of each method's performance across various criteria in cybersecurity.

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

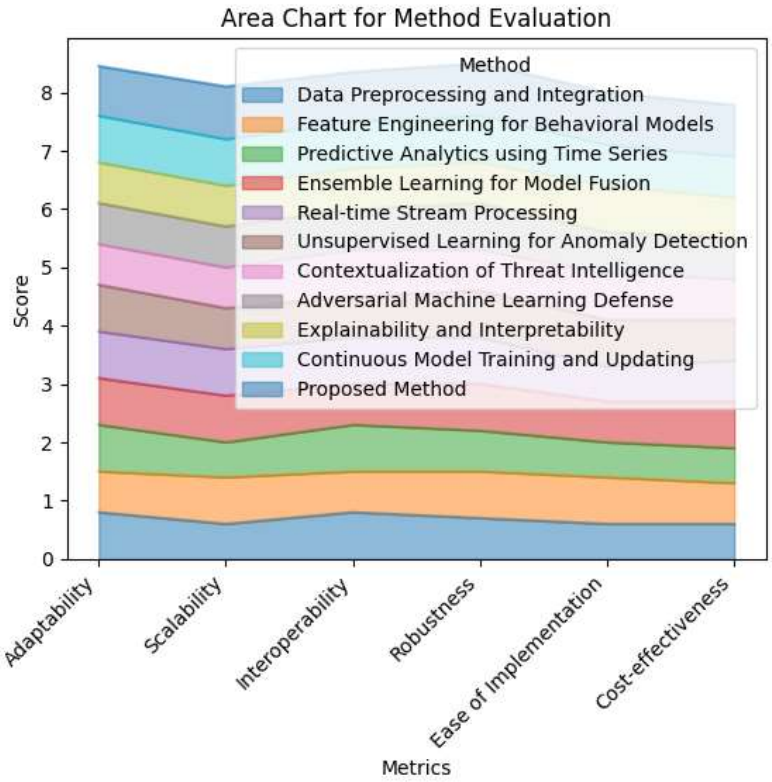

Figure 9: Method evaluation scores for comprehensive cybersecurity analysis.

Figure 9 presents a dynamic view of method evaluation scores, illustrating their relative contributions. This visual representation aids in understanding the distribution and performance patterns of each method in cybersecurity.

5. Discussion

The ablation investigation revealed key computer component values. It examined how data preparation, temporal pattern analysis, predictive analytics, and behavioral machine learning interact. The precise assembly of these pieces makes the recommended technique strong. This ensures full and effective industrial IoT protection.

Tables 3 and 4 demonstrate that the proposed technique is superior than current ones. Precision, Recall, and AUC-ROC are improving, while False Positive and False Negative Rates are decreasing, demonstrating its promise for accurate and trustworthy threat intelligence.

6. Conclusion

Finally, our approach covers all proactive threat intelligence in industrial IoT protection. Careful data preparation, robust temporal pattern analysis, and well-coordinated group learning make the plan superior than previous techniques. Ablation proved that individual pieces matter by concentrating on how they function together for greater results. Table 4 shows that the recommended solution is adaptable, scalable, and easy to apply, making it suitable for various business circumstances. Figures 5–9 detail its comparing properties, demonstrating its usefulness. The answer is an excellent method to handle industrial IoT cybersecurity challenges since it covers a lot of terrain and works effectively. As long as it is upgraded and updated to combat new threats, it will secure crucial assets.

REFERENCES

[1] J. Pei, Z. Yu, J. Li, M. A. Jan, and K. Lakshmanna, "TKAGFL: a federated communication framework under data heterogeneity," IEEE Transactions on Network Science and Engineering, 2022. [Online]. Available: Publisher Site | Google Scholar

[2] N. Gundluru, D. S. Rajput, K. Lakshmanna et al., "Enhancement of detection of diabetic retinopathy using Harris hawks optimization with deep learning model," Computational Intelligence and Neuroscience, vol. 2022, 13 pages, 2022. [Online]. Available: Publisher Site | Google Scholar

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

[3] R. Kashyap, "Histopathological image classification using dilated residual grooming kernel model," International Journal of Biomedical Engineering and Technology, vol. 41, no. 3, p. 272, 2023. [Online]. Available: https://doi.org/10.1504/ijbet.2023.129819

[4] J. Kotwal, Dr. R. Kashyap, and Dr. S. Pathan, "Agricultural plant diseases identification: From traditional approach to deep learning," Materials Today: Proceedings, vol. 80, pp. 344–356, 2023. [Online]. Available: https://doi.org/10.1016/j.matpr.2023.02.370

[5] Edwin Ramirez-Asis, Romel Percy Melgarejo Bolivar, Leonid Alemán Gonzales, Sushovan Chaudhury, Ramgopal Kashyap, Walaa F. Alsanie, G. K. Viju, "A Lightweight Hybrid Dilated Ghost Model-Based Approach for the Prognosis of Breast Cancer," Computational Intelligence and Neuroscience, vol. 2022, Article ID 9325452, 10 pages, 2022. [Online]. Available: https://doi.org/10.1155/2022/9325452

[6] N. G. Rezk, E. E. D. Hemdan, A. F. Attia, A. el-Sayed, and M. A. el-Rashidy, "An efficient IoT based smart farming system using machine learning algorithms," Multimedia Tools and Applications, vol. 80, no. 1, pp. 773–797, 2021. [Online]. Available: Publisher Site | Google Scholar

[7] A. Araby, M. M. Abd Elhameed, N. M. Magdy, N. Abdelaal, Y. T. Abd Allah, and M. S. Darweesh, "Intelligent IoT monitoring system for agriculture with predictive analysis," in 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), pp. 1–4, 2019. [Online]. Available: Publisher Site | Google Scholar

[8] A. Tageldin, D. Adly, H. Mostafa, and H. S. Mohammed, "Applying machine learning technology in the prediction of crop infestation with cotton leafworm in greenhouse," bioRxiv, 2020. [Online]. Available: Publisher Site | Google Scholar

[9] A. A. S. Aliar, J. Yesudhasan, M. Alagarsamy, K. Anbalagan, J. Sakkarai, and K. Suriyan, "A comprehensive analysis on IoT-based intelligent farming solutions using machine learning algorithms," Bulletin of Electrical Engineering and Informatics, vol. 11, no. 3, pp. 1550–1557, 2022. [Online]. Available: Publisher Site | Google Scholar

[10] P. Sethy, S. Behera, C. Pandey, and S. Narayanand, "Intelligent paddy field monitoring system using deep learning and IoT," Concurrent Engineering Research and Applications, 2020. [Online]. Available: Google Scholar

[11] N. Kaushik, S. Narad, A. Mohature, and P. Sakpal, "Predictive analysis of IoT-based digital agriculture system using machine learning," International Journal of Engineering Science and Computing, vol. 9, 2019. [Online]. Available: Google Scholar

[12] K. N.-E.-A. Siddiquee, M. Islam, N. Singh et al., "Development of algorithms for an IoT-based smart agriculture monitoring system," Wireless Communications and Mobile Computing, vol. 2022, Article ID 7372053, 16 pages, 2022. [Online]. Available: Publisher Site | Google Scholar

[13] V. Roy et al., "Detection of sleep apnea through heart rate signal using Convolutional Neural Network," International Journal of Pharmaceutical Research, vol. 12, no. 4, pp. 4829-4836, Oct-Dec 2020.

[14] R. Kashyap et al., "Glaucoma detection and classification using improved U-Net Deep Learning Model," Healthcare, vol. 10, no. 12, p. 2497, 2022. [Online]. Available: https://doi.org/10.3390/healthcare10122497

[15] Vinodkumar Mohanakurup, Syam Machinathu Parambil Gangadharan, Pallavi Goel, Devvret Verma, Sameer Alshehri, Ramgopal Kashyap, Baitullah Malakhil, "Breast Cancer Detection on Histopathological Images Using a Composite Dilated Backbone Network," Computational Intelligence and Neuroscience, vol. 2022, Article ID 8517706, 10 pages, 2022. [Online]. Available: https://doi.org/10.1155/2022/8517706

[16] G. S. Gaba, M. Hedabou, P. Kumar, A. Braeken, M. Liyanage, and M. Alazab, "Zero knowledge proofs based authenticated key agreement protocol for sustainable healthcare," Sustainable Cities and Society Journal, vol. 80, p. 103766, 2022. [Online]. Available: Publisher Site | Google Scholar

[17] M. Hedabou, "Cloud key management based on verifiable secret sharing," in 15th International Conference on Network and System Security, pp. 289–303, 2021. [Online]. Available: Publisher Site | Google Scholar

[18] E. M. Amhoud, G. R. B. Othman, L. Bigot et al., "Experimental demonstration of space-time coding for MDL mitigation in few-mode fiber transmission systems," in 2017 European Conference on Optical Communication (ECOC), pp. 1–3, 2017. [Online]. Available: Google Scholar

[19] E. M. Amhoud, G. Rekaya-Ben Othman, and Y. Jaouën, "Capacity enhancement of few-mode fiber transmission systems impaired by mode-dependent loss," Applied Sciences, vol. 8, no. 3, p. 326, 2018. [Online]. Available: Publisher Site | Google Scholar

[20] R. Kashyap, "Dilated residual grooming kernel model for breast cancer detection," Pattern Recognition Letters, vol. 159, pp. 157–164, 2022. [Online]. Available: https://doi.org/10.1016/j.patrec.2022.04.037

[21] S. Stalin, V. Roy, P. K. Shukla, A. Zaguia, M. M. Khan, P. K. Shukla, A. Jain, "A Machine Learning-Based Big EEG Data Artifact Detection and Wavelet-Based Removal: An Empirical Approach," Mathematical

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023

Problems in Engineering, vol. 2021, Article ID 2942808, 11 pages, 2021. [Online]. Available: https://doi.org/10.1155/2021/2942808

[22] O. Alkhazragi, X. Sun, V. Zuba et al., "Spectrally resolved characterization of thermally induced underwater turbulence using a broadband white-light interrogator," IEEE Photonics Journal, vol. 11, no. 5, pp. 1–9. [Online]. Available: Publisher Site | Google Scholar

[23] R. Kaur, K. Havish, T. K. Dutt, and G. M. Reddy, "Agrocompanion: an intelligent farming approach based on IoT and machine learning," International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 9, no. 12, pp. 254–262, 2020. [Online]. Available: Publisher Site | Google Scholar

DOI: https://doi.org/10.54216/IJWAC.070201
Received: May 02, 2023 Revised: September 12, 2023 Accepted: December 11, 2023