International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

AI-Driven Autonomous Cyber Defence System: A Review Study

Prof. Sadaf Zama[1], Dhammadeep Ganvir[2], Adharv S Aneesh[3], Harshal Bhure[4], Ankit Nagpure[5], Utkarsh Shamkuwar[6]

Department of Computer Science and Engineering (Cyber Security)
G. H. Raisoni College of Engineering and Management, Nagpur (GHRCEMN), India

---

Abstract – Digital systems today are strongly dependent on network connectivity, which also increases the risk of cyber attacks. Security teams often struggle to monitor large amounts of network data and respond quickly to suspicious activities. Intelligent technologies from Artificial Intelligence can help overcome this challenge by enabling systems to analyze behavior patterns automatically. An autonomous cyber defence system can observe network traffic, identify unusual actions, and react without waiting for human instructions. Learning techniques from Machine Learning allow the system to improve its detection ability over time by studying previous incidents. When abnormal behavior is discovered, the defence mechanism can limit access, block harmful communication, or isolate affected parts of the network. Such an approach supports faster reaction to threats and strengthens overall protection of digital infrastructure. The concept demonstrates how intelligent automation can assist organizations in handling modern cybersecurity risks more effectively.

Key Words: Artificial Intelligence, Cybersecurity, Machine Learning, Autonomous Cyber Defence, Threat Detection, Anomaly Detection.

1. INTRODUCTION

The rapid growth of digital technologies has transformed the way organizations store, process, and transmit information. However, this digital transformation has also increased the number of cyber threats targeting networks and information systems. Cyber attackers continuously develop new techniques to exploit vulnerabilities and gain unauthorized access to sensitive data. As a result, traditional cybersecurity tools such as firewalls and signature-based intrusion detection systems are often unable to detect advanced and evolving attacks.

Recent developments in **Artificial Intelligence** have opened new possibilities for improving cybersecurity systems. AI technologies can analyze large volumes of network data and identify patterns that may indicate suspicious behavior. By using intelligent algorithms, cybersecurity systems can detect threats more efficiently and respond to them automatically.

An AI-Driven Autonomous Cyber Defence System is designed to monitor network activities, detect abnormal patterns, and automatically initiate defensive actions without requiring constant human intervention. By combining techniques from **Machine Learning** with real-time monitoring, such systems can improve the speed and accuracy of cyber threat detection.

2. BACKGROUND OF CYBER THREATS

The rapid growth of digital infrastructure has resulted in an increase in various forms of cyber attacks. Attackers use sophisticated methods to exploit vulnerabilities in networks and systems.

Some of the most common cyber threats include:

- **Phishing** – Fraudulent attempts to obtain sensitive information such as passwords or financial data.
- **Ransomware** – Malware that encrypts user data and demands payment for its release.
- **Distributed Denial-of-Service (DDoS) attack** – Attackers flood a network with excessive traffic to disrupt services.

Traditional security solutions are often unable to effectively handle these evolving threats. As a result, organizations require intelligent systems that can detect and respond to cyber attacks in real time.

3. PROBLEM STATEMENT

Despite the development of numerous cybersecurity tools, several important challenges still exist.

First, modern networks generate extremely large volumes of data. Security analysts find it difficult to manually analyze such data in order to identify potential threats.

Second, cyber attackers frequently develop new attack techniques that bypass traditional signature-based detection systems. These unknown attacks can remain undetected for long periods.

Third, many security systems require constant human monitoring, which increases operational costs and slows response time during security incidents.

These challenges highlight the need for an intelligent and automated cyber defence system capable of continuously monitoring network activity and responding to threats without human intervention.

4. OBJECTIVES OF THE STUDY

The main objectives of this research include:

1. To design an AI-based autonomous cyber defence system.
2. To detect abnormal network behavior using intelligent algorithms.
3. To improve the accuracy and speed of cyber threat detection.
4. To reduce dependence on manual monitoring in cybersecurity operations.
5. To provide automated responses for preventing cyber attacks.

5. LITERATURE REVIEW

The application of artificial intelligence in cybersecurity has gained significant attention in recent years. Researchers have explored various AI techniques to improve the detection and prevention of cyber threats. These approaches aim to overcome the limitations of traditional security systems by enabling automated analysis and intelligent decision-making.

In **2023**, significant research efforts were directed toward improving Intrusion Detection Systems (IDS) using machine learning approaches. Studies demonstrated that ML-based IDS can effectively analyze large volumes of network traffic and detect anomalies by learning behavioral patterns, rather than relying on traditional rule-based or signature-based methods. This approach proved highly effective in identifying unknown and zero-day attacks, thereby improving overall system security.

Another important contribution in **2023** involved the integration of Machine Learning with Explainable Artificial Intelligence (XAI) and Large Language Models (LLMs). This hybrid approach enhanced threat detection capabilities while also improving the interpretability of AI decisions. As a result, cybersecurity analysts were able to better understand system outputs, leading to reduced false positives and increased trust in AI-driven security systems.

In **2024**, Mohammed M. S. and H. A. Talib conducted a comprehensive review of machine learning algorithms applied to intrusion detection systems. Their study highlighted that algorithms such as Random Forest, Support Vector Machines (SVM), and Neural Networks significantly improve detection accuracy. Furthermore, their findings emphasized the ability of these models to automate threat detection in large-scale and complex network environments.

The year **2025** witnessed major advancements in AI-driven cybersecurity systems. Research on machine learning-based IDS indicated that AI systems are capable of real-time threat detection by analyzing large-scale datasets and identifying abnormal behavioral patterns. However, these studies also identified critical challenges, including adversarial attacks, data privacy concerns, and issues related to model interpretability.

Additionally, several studies in **2025** proposed hybrid intrusion detection models that combine multiple machine learning and deep learning techniques, such as Random Forest, Long Short-Term Memory (LSTM) networks, and Graph Neural Networks (GNNs). These hybrid models demonstrated superior performance in detecting complex and multi-stage cyber attacks across distributed network environments.

In the same year, Naseem Khan et al. explored the application of Explainable AI (XAI) in intrusion detection systems. Their systematic review concluded that XAI enhances transparency, interpretability, and trustworthiness of AI-based cybersecurity solutions, making them more suitable for deployment in real-world and Industry 5.0 environments.

Furthermore, Ali Hozouri and his research team conducted a comprehensive survey in **2025** focusing on recent developments in machine learning and deep learning-based intrusion detection systems. Their study emphasized the importance of developing adaptive,

scalable, and intelligent security models capable of responding to continuously evolving cyber threats.

Another study in **2025** highlighted that AI-powered intrusion detection systems provide significant advantages over traditional methods, including real-time monitoring, faster anomaly detection, and reduced false positive rates. These improvements contribute to more efficient and reliable cybersecurity frameworks.

In **2026**, recent research trends indicate a strong emphasis on the integration of Explainable AI and advanced machine learning techniques in cybersecurity systems. Modern AI-based solutions are capable of processing massive volumes of network data, improving detection accuracy, and enabling automated decision-making. The growing importance of XAI is particularly notable, as it enhances system transparency, reliability, and user trust in AI-driven cybersecurity applications.

6. PROPOSED AI-DRIVEN CYBER DEFENCE SYSTEM

The proposed AI-Driven Autonomous Cyber Defence System is designed to provide intelligent and automated protection against modern cyber threats. Traditional cybersecurity solutions often rely on static rules and predefined signatures, which makes them ineffective against newly emerging attacks. To address these limitations, the proposed system integrates advanced techniques from **Artificial Intelligence**, **Machine Learning**, and network security analytics.

The primary objective of the system is to continuously monitor network activities, analyze behavioral patterns, detect malicious activities, and automatically initiate defensive actions without requiring constant human supervision. By combining real-time data analysis with intelligent decision-making, the system is capable of identifying complex cyber threats and responding to them rapidly.

6.1. System Overview

The proposed cyber defence system consists of multiple interconnected modules that work together to provide comprehensive security protection. These modules include the data collection module, data preprocessing module, AI-based analysis engine, threat detection module, automated response system, and monitoring dashboard.

The system continuously gathers network traffic information, system logs, and user activity data from different sources. This information is then processed and analyzed using intelligent algorithms to identify abnormal patterns that may indicate cyber attacks.

Once a suspicious activity is detected, the system automatically initiates appropriate security actions such as blocking malicious traffic, isolating compromised devices, or generating alerts for system administrators.

6.2. Data Collection and Monitoring

The first stage of the proposed system involves continuous monitoring of network activities. The data collection module captures various types of information, including:

- Network traffic data
- System log files
- User authentication records
- Application activity logs
- External threat intelligence data

This data is collected in real time from routers, firewalls, servers, and endpoint devices. The collected information provides valuable insights into network behavior and helps the system detect suspicious activities.

6.3. Data Preprocessing and Feature Extraction

Before the collected data can be analyzed, it must be preprocessed to remove unnecessary or irrelevant information. Data preprocessing improves the quality of the dataset and ensures that the AI algorithms can operate efficiently.

The preprocessing stage includes several important steps:

- Data cleaning and normalization
- Removal of duplicate records
- Feature extraction from network traffic
- Conversion of raw data into structured format

Feature extraction plays a critical role in this stage because it identifies important characteristics of network behavior that can help detect cyber attacks.

6.4. AI-Based Threat Detection Engine

The core component of the proposed system is the AI analysis engine. This module uses machine learning models to analyze network traffic and identify abnormal patterns that may indicate malicious activity.

Several machine learning algorithms can be used for threat detection, including:

- Decision Trees
- Support Vector Machines
- Random Forest Algorithms
- Neural Networks

These algorithms analyze large datasets and learn patterns associated with both normal and malicious network activities. Over time, the system improves its detection accuracy by continuously learning from new data.

Advanced deep learning techniques can also be applied to detect complex cyber threats that traditional security systems might miss.

6.5. Automated Threat Response Mechanism

One of the most important features of the proposed cyber defence system is its ability to automatically respond to detected threats. When suspicious behavior is identified, the automated response module immediately executes predefined security actions.

Examples of automated response actions include:

- Blocking malicious IP addresses
- Isolating compromised devices from the network
- Terminating suspicious processes
- Updating firewall rules
- Generating real-time alerts for administrators

This automated response significantly reduces the time required to mitigate cyber attacks and minimizes potential damage to the network.

6.6. Advanced Security Features

The proposed AI-Driven Autonomous Cyber Defence System integrates several advanced security capabilities that enhance its effectiveness in detecting and mitigating sophisticated cyber threats. These features enable the system to analyze complex network behaviors, predict potential attacks, and adapt its defensive strategies dynamically. By combining intelligent algorithms with real-time monitoring, the system provides a proactive cybersecurity framework capable of protecting modern digital infrastructures.

6.6.1. Behavioral Analysis

Behavioral analysis is a critical component of modern cybersecurity systems. Instead of relying solely on predefined attack signatures, the proposed system studies the **normal behavior patterns of users, applications, and network devices**.

The system collects behavioral data from several sources, including:

- System log files
- User authentication records
- Application activity logs
- Network traffic patterns

Using this information, machine learning models establish a baseline profile representing normal system behavior. For example, the system learns typical login times, frequently accessed resources, network bandwidth usage, and standard communication patterns between devices.

Once the baseline behavior is established, the system continuously compares real-time activity against the learned patterns. Any significant deviation from the expected behavior is flagged as a potential security threat. For instance:

- A user logging in from an unusual geographic location
- A sudden increase in network traffic from a specific device
- Unauthorized access attempts to restricted files

These anomalies may indicate insider threats, compromised accounts, or malicious activities. Behavioral analysis is particularly effective in detecting **insider attacks**, which are difficult to identify using traditional rule-based security systems.

Advanced anomaly detection algorithms allow the system to analyze both **temporal patterns** and **network**

relationships, enabling more accurate detection of suspicious behavior.

6.6.2. Real-Time Threat Intelligence

Real-time threat intelligence enhances the system's ability to detect known cyber threats quickly. Threat intelligence refers to information collected about existing cyber attacks, malicious software, and attacker behaviors.

The proposed system integrates external threat intelligence feeds obtained from cybersecurity research organizations and security databases. These sources provide continuously updated information such as:

- Lists of malicious IP addresses
- Known malware signatures
- Phishing domain databases
- Attack techniques used by cybercriminal groups

By integrating this intelligence with the system's monitoring infrastructure, the defence system can instantly recognize known threats.

For example, if network traffic originates from a suspicious IP address listed in the threat intelligence database, the system immediately blocks the connection and generates an alert.

Real-time threat intelligence also helps identify large-scale cyber attack campaigns. By correlating threat information with internal network data, the system can detect coordinated attacks targeting multiple systems simultaneously.This capability significantly improves the speed of threat detection and reduces the risk of successful cyber intrusions.

6.6.3. Self-Learning Capability

One of the most powerful features of AI-based cybersecurity systems is their ability to **continuously learn and improve**. The proposed system incorporates self-learning mechanisms based on machine learning algorithms.

Traditional security systems require manual updates whenever new cyber threats emerge. In contrast, AI-based systems can automatically update their knowledge by analyzing new data.

The self-learning process involves three major stages:

1.Data Collection: The system collects large volumes of historical network data, including both normal activities and previous cyber attack incidents.

2.Model Training: Machine learning models are trained using this dataset to distinguish between legitimate network behavior and malicious activities.

3.Continuous Learning: As new network events occur, the system updates its models to adapt to changing network conditions and evolving cyber threats.

This continuous learning process enables the system to detect previously unknown threats that may not exist in traditional signature databases. Additionally, advanced deep learning techniques allow the system to identify complex attack patterns across multiple network layers, further improving detection accuracy.

6.6.4. Predictive Threat Analysis

The system can analyze historical attack data and predict potential cyber threats before they occurPredictive threat analysis is an advanced capability that allows the system to anticipate potential cyber attacks before they occur.

The system uses historical security data, attack trends, and behavioral analytics to identify patterns that may indicate an upcoming attack. By applying predictive algorithms, the system can estimate the probability of certain cyber threats occurring within a specific time period.

For example, the system may detect an increasing number of suspicious login attempts targeting a specific server. By analyzing historical data, the system can predict that the server may soon be targeted by a brute-force attack.

Predictive analysis enables organizations to take **preventive actions**, such as:

- Strengthening firewall rules
- Increasing monitoring of critical systems
- Implementing additional authentication mechanisms

This proactive approach shifts cybersecurity from a **reactive defense strategy** to a **predictive security model**, significantly reducing the likelihood of successful cyber attacks.

6.6.5. Adaptive Security Mechanisms

Adaptive security refers to the system's ability to dynamically adjust its defence strategies based on the evolving threat environment.

Cyber attackers constantly develop new techniques to bypass security systems. Therefore, static security rules are often insufficient to protect modern networks. The proposed system continuously evaluates the effectiveness of its security policies and automatically modifies them when necessary. This adaptive capability is achieved through the integration of machine learning models with automated policy management.

For instance, if the system detects repeated attack attempts from a particular network region, it may automatically tighten firewall restrictions or implement additional authentication requirements for users accessing the system from that region.

Similarly, if abnormal traffic patterns are detected within a specific network segment, the system may temporarily isolate that segment until the threat is analyzed. Adaptive security mechanisms ensure that the defence system remains effective even as cyber threats evolve. This dynamic approach allows organizations to maintain strong security protection without requiring constant manual intervention.

7. SYSTEM ARCHITECTURE

The architecture of the proposed AI-driven autonomous cyber defence system is designed to provide a **multi-layered intelligent security framework** capable of detecting, analyzing, and mitigating cyber threats in real time. The system is composed of several interconnected layers that work together to collect security data, analyze patterns using intelligent algorithms, identify potential threats, and automatically respond to malicious activities.

Each architectural layer performs a specialized function while maintaining continuous communication with other modules in order to ensure efficient threat detection and response.

7.1. Data Collection Layer

The **Data Collection Layer** is responsible for gathering security-related information from multiple sources across the network infrastructure. This layer forms the foundation of the entire defence system because the accuracy of threat detection depends heavily on the quality and completeness of the collected data.

The system continuously collects information from various sources such as:

- Network traffic packets
- System event logs
- Firewall logs
- User authentication records
- Endpoint device activity
- Application usage logs

Network sensors, monitoring tools, and logging mechanisms are deployed across different nodes of the network to capture this data in real time. The collected information is stored in centralized repositories or security information management systems for further analysis.

The primary objective of this layer is to create a **comprehensive dataset representing normal and abnormal network behaviors**, which will later be analyzed by intelligent algorithms.

7.2. Data Processing Layer

Once the raw data is collected, it is forwarded to the **Data Processing Layer** for preparation and transformation. Raw network data often contains redundant, incomplete, or irrelevant information that may reduce the efficiency of AI-based detection systems.

Therefore, this layer performs several important preprocessing operations, including:

- Data cleaning and removal of corrupted records
- Data normalization and standardization
- Feature extraction from network traffic
- Removal of duplicate entries
- Data transformation into structured formats

Feature extraction is particularly important because it identifies key attributes that help distinguish between normal network behavior and malicious activity. Examples of extracted features may include packet size, communication frequency, login attempts, and network latency.

The processed data is then converted into a structured dataset that can be effectively analyzed by machine learning models.

7.3. AI Analysis Engine

The **AI Analysis Engine** represents the core intelligence of the proposed cyber defence system. This component utilizes algorithms from **Machine Learning** and advanced data analytics techniques to analyze network activity patterns and identify suspicious behaviors.

The AI engine is trained using historical cybersecurity datasets that contain both normal and malicious network activities. By learning from these datasets, the system can recognize complex attack patterns and detect unusual behaviors that may indicate cyber threats.

Several machine learning techniques may be applied within this module, including:

- Supervised learning for attack classification
- Unsupervised learning for anomaly detection
- Deep learning models for complex pattern recognition
- Clustering algorithms for identifying abnormal network segments

These models continuously analyze incoming data streams and compare them with previously learned patterns. If the system identifies significant deviations from expected behavior, the activity is flagged for further investigation.

The AI engine also improves over time through continuous learning, allowing it to adapt to new attack techniques and evolving cyber threats.

7.4. Threat Detection Module

The **Threat Detection Module** is responsible for identifying potential cyber attacks based on the analytical results produced by the AI engine.

This module evaluates the probability that a particular activity may represent a cyber threat. If suspicious behavior is detected, the system classifies the threat based on its severity and type. Examples of detected threats may include:

- Distributed denial-of-service (DDoS) attacks
- Malware infections
- Unauthorized access attempts
- Phishing attacks
- Insider threats

The detection module applies predefined risk scoring mechanisms to determine the seriousness of each detected event. High-risk threats trigger immediate defensive responses, while lower-risk activities may be monitored further for additional evidence.

This layered detection approach helps reduce false positives while ensuring that genuine cyber threats are identified quickly.

7.5. Automated Response Module

The **Automated Response Module** enables the cyber defence system to react to detected threats without requiring immediate human intervention. Once a cyber threat is confirmed, the system automatically initiates

appropriate security measures to minimize potential damage.

Some of the automated defensive actions include:

- Blocking malicious IP addresses
- Terminating suspicious network connections
- Isolating compromised devices from the network
- Updating firewall security rules

This module plays a crucial role in transforming the security system from a **reactive defence mechanism** into a **proactive and autonomous cyber defence solution.**

7.6. Reporting and Alert System

The **Reporting and Alert System** provides visibility and transparency into the security operations of the cyber defence system. This module generates alerts, logs security incidents, and provides analytical reports for system administrators.

When a potential cyber threat is detected, the system immediately sends notifications through various channels such as:

- Email alerts
- Security dashboards
- SMS notifications
- Incident management systems

The reporting system also maintains detailed logs of all detected security incidents. These logs can be used for further forensic investigation and to improve the system's detection capabilities.

Additionally, administrators can view visual reports and statistical summaries that illustrate attack trends, system vulnerabilities, and overall network security performance. These insights help organizations develop more effective cybersecurity strategies and improve their overall defence posture.

8. IMPLEMENTATION METHODOLOGY

The implementation of the proposed AI-Driven Autonomous Cyber Defence System follows a structured methodology consisting of several stages. Each stage plays an important role in ensuring that the system can accurately detect cyber threats and automatically respond to malicious activities. The implementation process involves data acquisition, preprocessing, model training, real-time threat detection, and automated response mechanisms.

8.1. Data Collection

The first stage in implementing the system involves collecting network and system data from multiple sources within the organizational infrastructure. Accurate and diverse data is essential for training machine learning models capable of identifying cyber threats.

Data is collected from sources such as:

- Network traffic packets captured using monitoring tools
- Firewall and intrusion detection system logs
- Server event logs
- User authentication records
- Application usage logs
- Endpoint device monitoring systems

These data sources provide valuable insights into the behavior of users, applications, and network devices. The collected data includes information such as IP addresses, packet sizes, communication frequency, login attempts, and session durations.

To ensure continuous monitoring, network sensors and log management tools are deployed throughout the network environment. The collected information is stored in a centralized data repository where it can be accessed for further analysis and model training.

8.2. Data Preprocessing

Raw network data is often unstructured and may contain redundant or incomplete information. Therefore, the second stage of the implementation involves preprocessing the collected data to improve its quality and suitability for machine learning analysis.

Data preprocessing includes several operations such as:

- Removing duplicate or corrupted records
- Handling missing values
- Normalizing numerical attributes
- Converting categorical data into numerical representations
- Extracting relevant features from network traffic

Feature extraction is particularly important because it identifies the most significant attributes that contribute to detecting malicious activities. Examples of useful features include packet transmission rate, number of login attempts, frequency of access to sensitive resources, and abnormal communication patterns.

After preprocessing, the cleaned dataset is organized into structured formats that can be used for training and testing machine learning models.

8.3. Model Training

In this stage, machine learning models are trained using historical cybersecurity datasets that contain examples of both normal network activities and cyber attacks. These datasets allow the system to learn patterns associated with malicious behavior.

Several machine learning algorithms can be applied during the training phase, including:

- Decision Tree classifiers
- Random Forest algorithms
- Support Vector Machines
- Neural networks for deep learning analysis

The training process involves feeding labeled data into the algorithms so that the models can learn to differentiate between legitimate network behavior and potential threats. The dataset is typically divided into training and testing subsets in order to evaluate the accuracy and performance of the trained model.

During this stage, the system also performs hyperparameter tuning to optimize the performance of the algorithms and reduce false positive rates.

8.4. Threat Detection

Once the machine learning models have been trained and validated, they are deployed within the cyber defence system to analyze real-time network traffic.

The threat detection process involves continuously monitoring network activity and comparing incoming data with the patterns learned during the training phase. If the system detects deviations from normal behavior, it flags the activity as a potential cyber threat.

Examples of detected anomalies may include:

- Unusual login attempts from unknown locations
- Rapid increases in network traffic
- Unauthorized access to restricted files
- Communication with suspicious external servers

The AI models assign risk scores to detected anomalies based on the probability that they represent malicious activities. If the risk score exceeds a predefined threshold, the system categorizes the event as a cyber attack. This intelligent detection mechanism allows the system to identify both known and previously unknown threats.

8.5. Automated Response

The final stage of the implementation methodology involves responding to detected threats through automated security actions. Once the system confirms the presence of a cyber attack, the automated response module immediately initiates defensive measures.

Possible automated responses include:

- Blocking malicious IP addresses
- Terminating suspicious network sessions
- Isolating compromised devices from the network
- Updating firewall and access control policies
- Alerting system administrators about the incident
-

The automated response mechanism significantly reduces the time required to mitigate cyber attacks and prevents attackers from spreading further within the network.

8.6. Continuous System Improvement

After deployment, the cyber defence system continues to improve its detection capabilities through continuous learning. New security incidents are incorporated into the dataset and used to retrain machine learning models.

This adaptive learning approach ensures that the system remains effective against evolving cyber threats and maintains high detection accuracy over time.

9. FUTURE WORK

Future research can focus on enhancing the accuracy, scalability, and adaptability of AI-driven cyber defence systems. With the rapid evolution of cyber threats, traditional machine learning models may not always be

sufficient to detect highly sophisticated attacks. Therefore, advanced approaches such as deep neural networks and reinforcement learning can be explored to develop more intelligent and adaptive security mechanisms. These techniques can enable the system to learn complex attack patterns, improve anomaly detection capabilities, and reduce false positive rates. Additionally, incorporating continuous learning frameworks can allow the defence system to update its knowledge automatically as new types of cyber threats emerge.

Another important direction for future research involves the integration of AI-based cyber defence systems with emerging technologies such as cloud computing and Internet of Things (IoT) environments. As modern digital infrastructures increasingly rely on interconnected devices and distributed networks, securing these systems becomes more challenging. Future work may focus on designing lightweight and scalable security models that can protect IoT devices and cloud platforms in real time. Furthermore, combining AI-based threat intelligence with distributed security architectures may significantly enhance the resilience of modern networks against large-scale cyber attacks.

10. CONCLUSION

1) This research paper proposes an **AI-Driven Autonomous Cyber Defence System** aimed at improving the detection and prevention of modern cyber threats.
2) Traditional cybersecurity systems mainly rely on signature-based detection methods, which are often ineffective against newly emerging and sophisticated cyber attacks.
3) The proposed system utilizes techniques from **Machine Learning** to analyze network behavior and identify abnormal activities in real time.
4) The architecture of the system includes several key components such as data collection, data preprocessing, AI analysis engine, threat detection module, automated response module, and reporting system.
5) Advanced security features such as behavioral analysis, predictive threat detection, and adaptive defence mechanisms enhance the efficiency and reliability of the system.
6) The automated response capability allows the system to quickly react to cyber threats by blocking malicious traffic and isolating compromised network components.
7) The proposed AI-driven cybersecurity framework improves detection accuracy, reduces response time, and provides proactive protection against evolving cyber threats.
8) Overall, the implementation of intelligent cyber defence systems can significantly strengthen the security of modern digital infrastructures and help organizations protect sensitive information from cyber attacks.

11. REFERENCES

[1] A. Kumar, R. Singh, and P. Sharma, "Explainable Artificial Intelligence for Cybersecurity: Enhancing Transparency in Intrusion Detection Systems ,"*International Journal of Computer Science and Information Security*, vol. 22, no. 1, pp. 45–58, **2026**.

[2]N.Khan, M. Ali, and S. Rahman, "Explainable AI-Based Intrusion Detection Systems for Industry 5.0Applications, "*Information*, vol. 16, no. 12, pp. 1–20, **2025**.

[3]A. Hozouri, M. Dehghani, and H. Karimipour, "Recent Advances in Machine Learning and Deep Learning for Intrusion Detection Systems,"*Journal of Cybersecurity and Privacy*, vol. 5, no. 2, pp. 210–230, **2025**.

[4] S. Verma and K. Patel,"Hybrid Machine Learning Models for Network Intrusion Detection Using LSTM and Random Forest,"*IEEE Access*, vol. 13, pp. 10234–10250, **2025**.

[5] R. Gupta and P. Mehta, "Machine Learning-Based Intrusion Detection Systems for Real-Time Cyber Threat Detection," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 4, pp. 120–130, **2025**.

[6] M. M. S. Mohammed and H. A. Talib,"A Review on Machine Learning Algorithms in Intrusion Detection Systems ," *International Journal of Scientific Research in Computer Science Engineering*, vol. 12, no. 1, pp. 75–85, **2024**.