Open Access International Journal of Science and Engineering

ISSN: 2456-3293

Volume 9 Issue 6 (2026) | Pages: 20-24

Machine Learning–Driven Analytical Framework for Detecting Cybersecurity Threats and System Vulnerabilities

Pooja Kamble[1], Mansi Bhate[2]

[1]ATSS CBSCA, Chinchwad, Pune
[2]IMCC, Kothrud, Pune, India

Peer Review Information

***Type**: Article*
***Received**: 26 March 2026*
***Revised**: 20 April 2026*
***Accepted**: 28 May 2026*
***Published**: 17 June 2026*

Abstract

Organisations are becoming far more vulnerable to cybersecurity risks due to the quick development of digital technology, cloud computing, and networked systems. Traditional security measures, such as intrusion detection systems based on signatures and rule-based firewalls are no longer adequate to deal with complex and changing threat patterns.

Examining a large amount of network data and identifying patterns linked to negative events can be accomplished with machine learning. What we do is not the same as machine learning. Models for machine learning can appear at what happened in the past and use that to find things that're not normal. This helps us find risks that we know about and ones that we do not know about with machine learning. A machine learning-based analytical approach for identifying cybersecurity risks and vulnerabilities in network environments is presented in this study. The suggested system incorporates several phases, such as feature engineering, preprocessing, data collecting, and classification through algorithms for supervised learning. Evaluation of experiments employing Benchmark datasets show more detection accuracy and fewer false alarm rates as compared to conventional systems.

The presented framework helps formulate intelligent, proactive, and flexible cyber security systems, which can address new risks in modern digital infrastructure.

Keywords: Cybersecurity; Machine Learning; Intrusion Detection; Threat Detection; Data Analytics; Network Security.

How to Cite This Article

Kamble, P., & Bhate, M. (2026). Machine Learning–Driven Analytical Framework for Detecting Cybersecurity Threats and System Vulnerabilities. *Open Access International Journal of Science and Engineering*, 9(6), 20–24.

Introduction

The way we manage data and talk to each other has really changed because of platforms. We are using things like cloud computing and the IOT to make things work better and faster. These new technologies are really good at handling a lot of work. They also help us connect lots of devices to the internet. The bad thing is that the latest technology gives cybercriminals more ways to attack us. Cybercriminals can now try to hurt us in more ways. Digital platforms and the IOT are making our lives easier. Also making it easier for cybercriminals to cause trouble.

Malware, phishing, distributed denial-of-service (DDoS), and unauthorised access are examples of contemporary cyber-attacks that have grown more sophisticated and challenging to identify. Conventional security solutions are ineffectual against new or developing threats because they rely on predetermined rules and known attack signatures.

By allowing computers to identify anomalies in real time and discover patterns from historical data, machine learning offers a data-driven solution. By recognising departures from the normal behaviour, ML models are better able to identify possible security issues.

By increasing accuracy and decreasing false positives in network security systems, this research seeks to provide a machine learning-based framework that improves intrusion detection capabilities.

Problem Statement

The current state of cyber security is marked by enormous amounts of network data and growing complexity. Monitoring such data manually is ineffective and frequently unfeasible. Furthermore, attackers are always changing how they get around conventional security measures. Static rules and pre-established attack fingerprints constitute the key components of traditional intrusion detection systems. Therefore, they cannot detect unknown vulnerabilities and zero-day attacks.

Another big worry is that these systems produce a lot of alarms. This creates work, for security analysts and makes it take longer to respond. So, we need a system that can automatically look at lots of data. It should be able to find patterns that're hard to see and accurately spot fake behaviour. Because machine learning can manage massive amounts of data, it can be used to address these problems.

Objectives

Below are the main objectives of this study:

- To analyse network traffic data in the identification of cybersecurity threats
- To design a machine learning model-based intrusion detection model
- To enhance detection accuracy through the use of predictive analytics
- To reduce false positive fractions in detection systems.
- To develop scalable cybersecurity analytics framework

Literature Review

There have been lot of studies about using machine learning in security. Sommer and Paxson talked about the problems of using machine learning in systems that detect intrusions. They said it is valuable to have datasets and to choose the right features for machine learning in cybersecurity.

Some people like Buczak and Guven did a lot of research on the methods of machine learning used in security. They observe that some algorithms like Random Forest and Support Vector Machines, and Neural Networks, can really help find cyber-attacks.

Kim and some other people suggested a system that can give intrusions in two ways. This system uses anomaly detection and misuse detection to make it work better.

To work this system smoothly, Kim and some other people said we should use a system that combines anomaly detection and misuse detection. By looking at a lot of network data we can use learning to find complicated patterns of cyber-attacks.

When we compare machine learning to the rule-based systems, machine learning is much better at detecting cyber-attacks, according to what users are saying now about machine learning, in cybersecurity.

Cybersecurity Threat Landscape

Cyber dangers are constantly changing and present serious risks to businesses. Typical threats consist of:

- Malware Attacks: Malicious software intended to steal data or interfere with systems
- Phishing Attacks: Dishonest attempts to obtain private data

- DDoS Attacks: Excessive traffic overloading systems
- Zero-Day Vulnerabilities: Taking advantage of unidentified system defects
- Insider Threats: When authorised users abuse their access rights

By analysing behavioural patterns, machine learning systems can identify abnormalities related to these dangers.

Proposed System Architecture

The proposed system consists of multiple components working together to detect cybersecurity threats.

Key Modules:

1. Data Collection: Gathers network traffic and security logs
2. Data Preprocessing: Cleans and prepares data
3. Feature Engineering: Extracts meaningful attributes
4. Model Training: Applies machine learning algorithms
5. Threat Detection Engine: Identifies malicious activities
6. Alert System: Generates alerts for suspicious behaviour
7. Dashboard: Provides visualization for monitoring

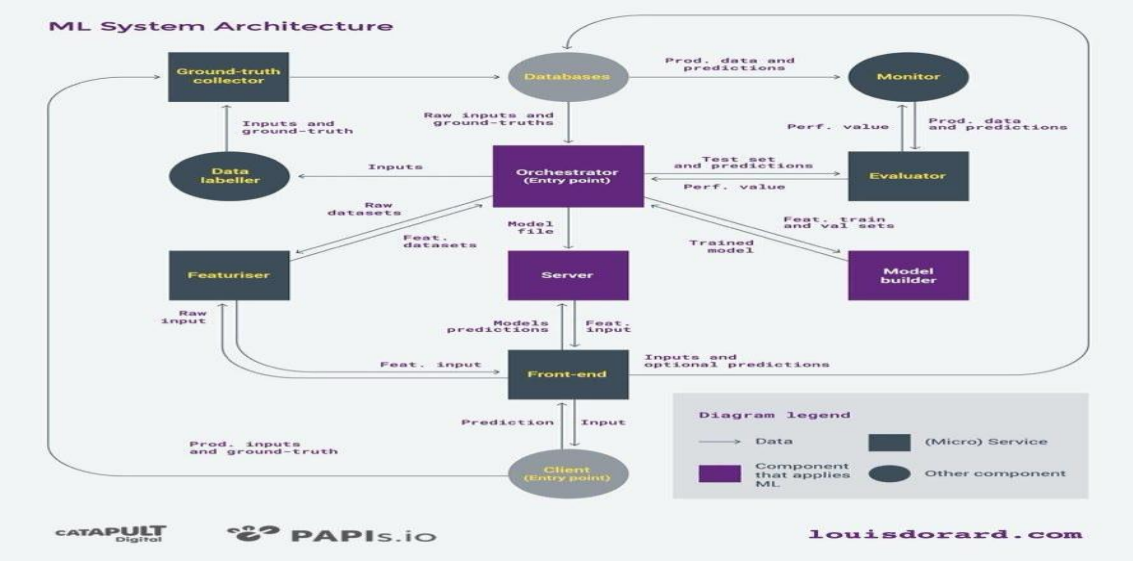

Fig. 1. *Architecture of a Machine Learning-Based Predictive System*

Techniques

Gathering Datasets

The research makes use of common databases like:

- KDD Cup 99
- UNSW-NB15
- NSL-KDD
- CICIDS2017

For training and testing, these datasets provide labelled network traffic data.

Data Pre-processing

The pre-processing stage includes:

- Handling missing values
- Encoding categorical variables
- Normalizing data
- Balancing datasets

Feature Selection

Important features include:

- Protocol type
- Packet size
- Connection duration
- Login attempts
- Network behavior

Machine Learning Algorithms

The following algorithms are used:

- Decision Tree: Rule-based classification
- Random Forest: Ensemble learning for higher accuracy
- Support Vector Machine: Optimal decision boundary detection
- Deep learning for intricate pattern recognition using NN

Implementation

Python and ML libraries are used in the system's implementation.

```
import pandas as pd
    from sklearn.model_selection import train_test_split
    from sklearn.ensemble import RandomForestClassifier
    from sklearn.metrics import accuracy_score
    data = pd.read_csv("dataset.csv")
    X = data.drop("label", axis=1)
    y = data["label"]
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
    model = RandomForestClassifier()
    model.fit(X_train, y_train)
    predictions = model.predict(X_test)
print ("Accuracy:", accuracy_score(y_test, predictions))
```

Performance Evaluation

The system is evaluated using:

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix

Results and Discussion

The experimental results indicate that ML models significantly improve threat detection accuracy. Because Random Forest and Neural Networks can identify intricate patterns in data, they perform better. The suggested methodology improves detection efficiency and lowers false positives as compared to conventional methods.

Advantages

- Detects unknown threats
- Automates data analysis
- Improves detection accuracy

- Enables real-time monitoring
- Reduces manual effort

Limitations

- Requires large datasets
- Computationally expensive
- Risk of over fitting
- Needs periodic updates

Prospects

Future projects could consist of:

- Combining deep learning models
- AI-based real-time detection systems
- Integration of cloud security
- Blockchain-driven security
- Federated learning strategies

Conclusion

Cyber security risks are increasing complicated because digital systems are expanding rapidly. Modern attacks are too powerful for security measures. This study showed a way to find weaknesses and cyber security risks using machine learning. The model uses ML techniques picking the features and preparing data to make detection better. The results say that machine learning makes it more accurate to detect intrusions and reduces wrong alarms. The suggested plan helps create flexible cyber security solutions that can protect modern digital systems. Machine learning and cyber security are really important for keeping systems safe, from cyber security risks.

References

1. L. Buczak and E. Guven, “A survey of machine learning methods for cybersecurity,” IEEE, 2016.
2. R. Sommer and V. Paxson, “Machine learning in intrusion detection,” IEEE, 2010.
3. G. Kim et al., “Hybrid intrusion detection system,” Elsevier, 2018.
4. I. Goodfellow et al., Deep Learning, MIT Press, 2016.
5. T. Tavallaee et al., “KDD dataset analysis,” IEEE, 2009.
6. M. Tavallaee et al., “NSL-KDD dataset,” 2009.
7. I. Sharafaldin et al., “CICIDS2017 dataset,” 2017.
8. N. Moustafa, “UNSW-NB15 dataset,” 2015.
9. J. Han et al., Data Mining Concepts, 2012.
10. S. Russell and P. Norvig, Artificial Intelligence, 2019.
11. W. Stallings, Network Security Essentials, 2017.
12. Bishop, Pattern Recognition, 2006.
13. J. Andress, Information Security Basics, 2014.
14. M. Bishop, Computer Security, 2018.
15. L. Breiman, “Random Forests,” 2001