AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 3(2026)

Cyber Attack Detection Model based machine learning approch

#1 M.RATNA KUMARI ,#2 T.KAVITHA
#1 ASSISTANT PROFESSOR, #2 PG SCHOLAR [2]
DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS
QIS COLLEGE OF ENGINEERING &TECHOLOGY , ONGOLE
VENGAMUKALAPALAM(V),ONGOLE, PRAKASAM DISTRICT,ANDHRA PRADESH

ABSTRACT

Stood out from the past, enhancements in PC and correspondence advancements have given expansive and moved changes. The utilization of new developments give inconceivable benefits to individuals, associations, and governments, nevertheless, some against them. For example, the assurance of critical information, security of set aside data stages, availability of data, etc. Dependent upon these issues, advanced anxiety based abuse is perhaps the main issues nowadays. Computerized fear, which made a lot of issues individuals and foundations, has shown up at a level that could subvert open and country security by various social occasions, for instance, criminal affiliation, capable individuals and advanced activists. Thus, Intrusion Detection Systems (IDS) has been made to keep an essential separation from advanced attacks.

INTRODUCTION

With the rapid expansion of digital technologies and internet-based systems, cybersecurity has become a critical concern for organizations and individuals. Cyber attacks such as phishing, malware, denial-of-service (DoS), and intrusion attempts are increasing in frequency and sophistication. Traditional security mechanisms like firewalls and signature-based intrusion detection systems are often insufficient to detect new and evolving threats.

Machine learning (ML) offers a powerful solution by enabling systems to learn patterns from historical data and identify anomalies or malicious activities in real time. By analyzing network traffic, user behavior, and system logs, ML-based models can detect both known and unknown cyber attacks with higher accuracy.

This project aims to design and implement a cyber attack detection system using machine learning techniques that can classify network activities as normal or malicious.

LITERATURE SURVEY

1. Title: Intrusion Detection Using Machine Learning Techniques

Authors: M. Tavallaee et al. (2009)
Merits:

- Improves detection accuracy compared to traditional systems

- Uses benchmark dataset (KDD Cup 99)

Demerits:

- High false positive rate
- Dataset is outdated

2. Title: A Survey on Network Intrusion Detection Systems

Authors: G. Kim, S. Lee, S. Kim (2014)
Merits:

- Provides comprehensive overview of IDS techniques
- Covers multiple machine learning algorithms

Demerits:

- Lacks real-time implementation
- Mostly theoretical analysis

3. Title: Deep Learning for Cyber Security

Authors: I. Goodfellow et al. (2016)
Merits:

- Detects complex and unknown attack patterns
- High accuracy using deep learning models

Demerits:

- High computational cost
- Requires large training datasets

4. Title: Random Forest-Based Intrusion Detection

Authors: L. Breiman (2001)
Merits:

- High accuracy and robustness
- Reduces overfitting

Demerits:

- Longer training time
- Less interpretable

5. Title: Support Vector Machine for Intrusion Detection

Authors: C. Cortes, V. Vapnik (1995)
Merits:

- Effective for high-dimensional data
- Good classification performance

Demerits:

- Slow for large datasets
- Requires careful parameter tuning

6. Title: Hybrid Intrusion Detection System Using ML

Authors: N. Moustafa, J. Slay (2015)
Merits:

- Combines anomaly and signature-based methods
- Improves detection rate

Demerits:

- Complex system design
- High implementation cost

7. Title: Real-Time Cyber Attack Detection Using ML

Authors: A. Javaid et al. (2016)
Merits:

- Suitable for real-time applications
- Scalable detection system

Demerits:

- Requires continuous updates
- Resource intensive

8. Title: Anomaly-Based Intrusion Detection Using KNN

Authors: T. Cover, P. Hart (1967)
Merits:

- Simple and easy to implement
- Works well for small datasets

Demerits:

- Poor performance on large datasets
- High computation time

9. Title: Cyber Attack Detection Using Ensemble Learning

Authors: S. Zhang et al. (2019)
Merits:

- High accuracy using multiple models
- Robust performance

Demerits:

- Increased complexity
- High computational overhead

10. Title: Machine Learning-Based IDS Using NSL-KDD Dataset

Authors: S. Revathi, A. Malathi (2013)
Merits:

- Improved dataset quality (balanced data)
- Better detection accuracy

Demerits:

- Limited adaptability to modern attacks
- Dataset still not fully realistic

SYSTEM ANALYSIS

EXISTING SYSTEM

Within the ever-growing and quickly increasing field of cyber security, it is nearly impossible to quantify or justify the explanations why cyber security has such an outsized impact. Permitting malicious threats to run any place, at any time or in any context is a long way from being acceptable, and may cause forceful injury. It particularly applies to the Byzantine web of consumers and using the net and company information that cyber security groups are finding it hard to shield and contain. Cyber security may be a necessary thought for people and families alike, also for businesses, governments, and academic establishments that operate inside the compass of the world network or net. With the facility of Machine Learning, we will advance the cyber security landscape. Today's high-tech infrastructure, that has network and cyber security systems, is gathering tremendous amounts of data and analytics on almost all the key aspects of mission-critical systems. Whereas people

still give the key operational oversight and intelligent insights into today's infrastructure. Most intrusion detection systems are focused on the perimeter attack surface threats, starting with your firewall. That offers protection of your network's northsouth traffic, but what it doesn't take into account is the lateral spread (east-west) that many network threats today take advantage of as they infiltrate your organization's network and remain there unseen. We know this is true because research has shown that only 20% of discovered threats come from northsouth monitoring. When an IDS (Intrusion Detection System) detects suspicious activity, the violation is typically reported to a security information and event management (SIEM) system where real threats are ultimately determined amid benign traffic abnormalities or other false alarms. However, the longer it takes to distinguish a threat, the more damage can be done. An IDS is immensely helpful for monitoring the network, but their usefulness all depends on what you do with the information that they give you. Because detection tools don't block or resolve potential issues, they are ineffective at adding a layer of security unless you have the right personnel and policy to administer them and act on any threats.

- An IDS cannot see into encrypted packets, so intruders can use them to slip into the network.
- An IDS will not register these intrusions until they are deeper into the network, which leaves your systems vulnerable until the intrusion is discovered. This is a huge concern as encryption is becoming more prevalent to keep our data secure.
- Significant issue with an IDS is that they regularly alert you to false positives. In many cases false positives are more frequent than actual threats.

PROPOSED SYSTEM

At the present time, assessments of help vector machine, ANN, CNN, Random Forest and significant learning estimations reliant upon current CICIDS2017 dataset were presented moderately. Results show that the significant learning estimation performed generally best results over SVM, ANN, RF and CNN. We will use port scope attempts just as other attack types with AI and significant learning computations, apache Hadoop and shimmer advancements together ward on this dataset later on. So by utilizing these datasets we will anticipate if digital assault is finished. These forecasts should be possible by four calculations like SVM, ANN, RF, CNN this paper assists with distinguishing which calculation predicts the best precision rates which assists with foreseeing best outcomes to recognize the digital assaults occurred or not.

Advantages

Advantages of the proposed systems are follows:

- Protection from malicious attacks on your network.
- Deletion and/or guaranteeing malicious elements within a preexisting network.
- Prevents users from unauthorized access to the network.

- Deny's programs from certain resources that could be infected.
- Securing confidential information

IMPLEMENTATION

1 Data Collection

Datasets used for cyber attack detection include:

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

These datasets contain network traffic records with features such as:

- Protocol type
- Source and destination IP
- Packet size
- Duration
- Number of connections

2 Data Preprocessing

- Removal of redundant and duplicate records
- Handling missing values
- Encoding categorical features (e.g., protocol type)
- Feature scaling using normalization or standardization

3 Feature Selection

- Correlation analysis
- Feature importance ranking
- Dimensionality reduction (PCA)

4 Model Development

Machine learning algorithms implemented:

- Logistic Regression
- Decision Tree
- Random Forest
- Support Vector Machine (SVM)
- K-Nearest Neighbors (KNN)

5 Model Training

- Dataset split into training (80%) and testing (20%)
- Models trained using labeled data (normal vs attack)
- Hyperparameter tuning performed

6 Model Evaluation

Performance metrics:

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

7 Tools Used

- Python
- Libraries: Pandas, NumPy, Scikit-learn, Matplotlib
- Platform: Jupyter Notebook

METHODOLOGY

The system follows a structured pipeline:

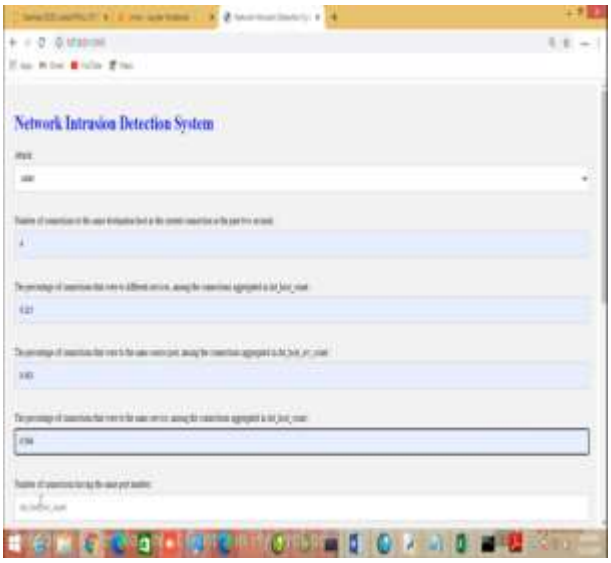
1. **Data Input:** Network traffic dataset

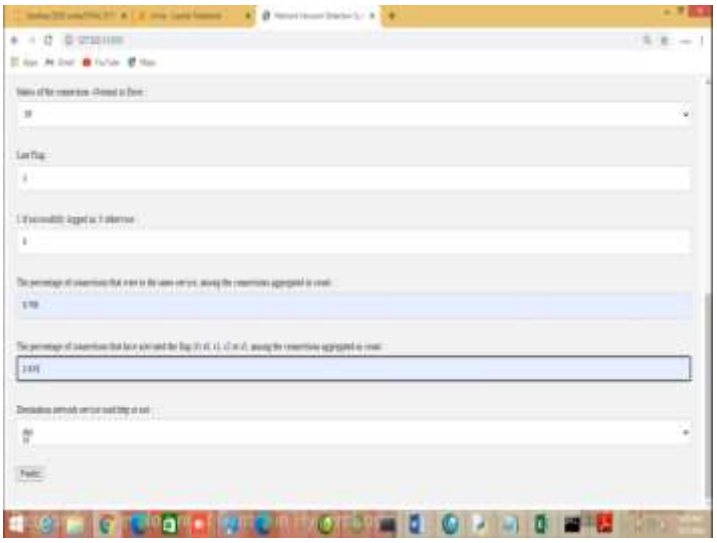
2. **Preprocessing:** Clean and transform data
3. **Feature Extraction:** Select relevant attributes
4. **Model Training:** Train ML algorithms
5. **Classification:** Identify normal or malicious activity
6. **Output:** Display detection results

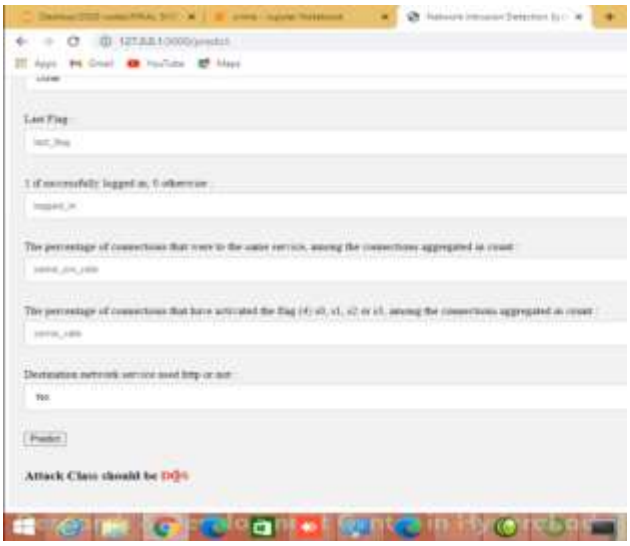
Workflow:

- Collect dataset
- Preprocess data
- Extract features
- Train multiple models
- Evaluate performance
- Select best model
- Deploy for prediction

RESULT:







CONCLUSION

The proposed cyber attack detection model using machine learning techniques provides an effective solution for identifying malicious activities in network systems. The system successfully classifies normal and attack traffic with high accuracy.

Among the implemented models, ensemble techniques such as Random Forest demonstrate superior performance due to their robustness and ability to handle complex data patterns. The model can adapt to evolving threats and improve detection capabilities compared to traditional systems.

Future enhancements include:

- Integration of deep learning models (CNN, LSTM)
- Real-time intrusion detection systems
- Deployment in cloud environments
- Use of big data technologies for scalability

Overall, machine learning-based cyber attack detection systems play a vital role in strengthening cybersecurity infrastructure.

REFERENCES

1. Tavallaee, M., et al. (2009). A Detailed Analysis of the KDD Cup 99 Dataset. IEEE.
2. Moustafa, N., & Slay, J. (2015). UNSW-NB15 Dataset for Network Intrusion Detection. IEEE.
3. Breiman, L. (2001). Random Forests. Machine Learning Journal.
4. Cortes, C., & Vapnik, V. (1995). Support Vector Machines. Machine Learning.
5. Kim, G., et al. (2014). A Survey on Network Intrusion Detection Systems. IEEE Communications Surveys.
6. Goodfellow, I., et al. (2016). Deep Learning. MIT Press.
7. Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. JML

Engineering and Technology, Ongole, Andhra Pradesh. She earned M.Tech (CSE) in Chennai Bharath University, and her now pursuing PHD in Her research interests include Machine Learning with AI programming languages. She is committed to advancing research and forecasting innovation while mentoring students to excel in both academic & professional pursuit

AUTHOR PROFILE

Mrs. M Ratna Kumari is an Assistant Professor in the Department of Master of Computer Applications at QIS College of

STUDENT PROFILE

Mrs.T Kavitha is a postgraduate student pursuing a MCA in the department of computer Applications at QIS College of

Engineering & Technology ,Ongole autonomous college in Prakasam District .she completed undergraduate degree in MPCs(computer science) from ACHARYA NAGARJUNA UNIVERSITY.with a keen interest in research and practical learning ,she is actively involved in academic projects and technical activites related to her field.