International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Network Intrusion Detection Using Supervised Machine Learning Technique with Feature Selection

Brugumalla Mahendra Achari[1], Mooramreddy Sreedevi[2]

[1]MCA Final Semester Student, Master of Computer Applications, SVU College of CM & CS, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

[2]Associate Professor, Master of Computer Applications, SVU College of CM & CS, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Abstract: ***In the increasingly digital world, ensuring the security of computer networks has become a crucial task. This paper introduces a machine learning-based solution to detect potential threats in network activity. By comparing the performance of two supervised learning models-Artificial Neural Networks (ANN) and Support Vector Machines (SVM)-alongside feature selection techniques, we found that ANN, when integrated with wrapper-based feature selection, yielded the highest accuracy on the NSL-KDD dataset. Our findings support the effectiveness of machine learning methods, particularly with refined feature inputs, in building robust and adaptable Network Intrusion Detection Systems (NIDS).***

Keywords: Cybersecurity, Machine Learning, Intrusion Detection, Anomaly Detection, ANN, SVM, Network Security, Feature Selection, Real-time Monitoring, False Positives, Attack Patterns

1. Introduction

With the expansion of internet use across the globe, the frequency and sophistication of cyberattacks have grown in parallel. Intrusion Detection Systems (IDS) serve as a vital tool to detect and prevent unauthorized access or threats to network systems. This study focuses on applying supervised machine learning algorithms—specifically SVM and ANN—to determine whether network traffic is benign or malicious. Given the growing dependence on digital infrastructure, it is imperative to secure systems against such threats effectively.

1.1 Existing System

Traditional intrusion detection systems, especially those using signature-based techniques, have limitations. While they are efficient at identifying known threats, they struggle to detect unfamiliar or zero-day attacks. Anomaly-based systems, which monitor deviations from expected behavior, offer better chances of identifying new threats, but often suffer from higher false positive rates. Our approach seeks to mitigate these issues through supervised learning.

1.2 Proposed System

Our work proposes a supervised machine learning model that learns from past network behavior to detect unseen attacks. We evaluated both SVM and ANN classifiers, ultimately finding that the ANN model, especially when used with a wrapper-based feature selection method, performed best. This model offers:

- Enhanced detection accuracy
- Faster response times using Hidden Naïve Bayes
- The ability to process large and complex traffic data

1.2.1 Advantages

- We introduced an enhanced model using hidden naïve bayes improving speed and accuracy
- It process large volumes of network data and adapts to complex attack behaviors

2. Literature Review

Modern information security depends on ensuring the confidentiality, integrity, and availability of data. IDS are designed to detect either known threats (via misuse detection) or anomalies. Combining both approaches (hybrid systems) often leads to more reliable detection.
Prior studies include:

- **Song et al. (2016)**: Examined cyber-victimization patterns.
- **Alaei & Noorbehbahani (2017)**: Developed an incremental anomaly-based IDS using limited labeled data.

3. Methodology

The proposed system employs:

- An intelligent attribute selection algorithm
- Improved Multiclass SVM (IREMSVM)

It learns from data and evolves through pattern recognition and rule-based decision-making. This intelligent, adaptive design leads to significantly better accuracy.

System Input and Output

- Input: Raw network traffic data (from NSL-KDD dataset)
- Output: Prediction of whether traffic is normal or indicates an attack

Key Components:

- Forward Propagation: Processes input data using weight multipliers
- Activation (Sigmoid) Function: Normalizes the output
- Back Propagation: Refines weights based on prediction errors to improve accuracy

Input 1	Input 2	Input 3	Output
0	1	1	1
1	0	0	0
1	0	1	1

Figure 2: Training Examples

Now we want to predict the output the following set of Inputs

1	0	1	?

Figure 3: Test Example

4. Results

Using the NSL-KDD dataset, the ANN model achieved a **detection accuracy of 96.88%**, outperforming SVM. The interface includes:

- Dataset upload
- Attack prediction
- Graphical representation of performance

Observation: ANN delivered better performance in terms of prediction accuracy than SVM, demonstrating its suitability for this application.

1) Forward Propagation:
Take the inputs, multiply by the weights (just use random numbers as weights)

Let $Y = W_iI_i = W_1I_1+W_2I_2+W_3I_3$

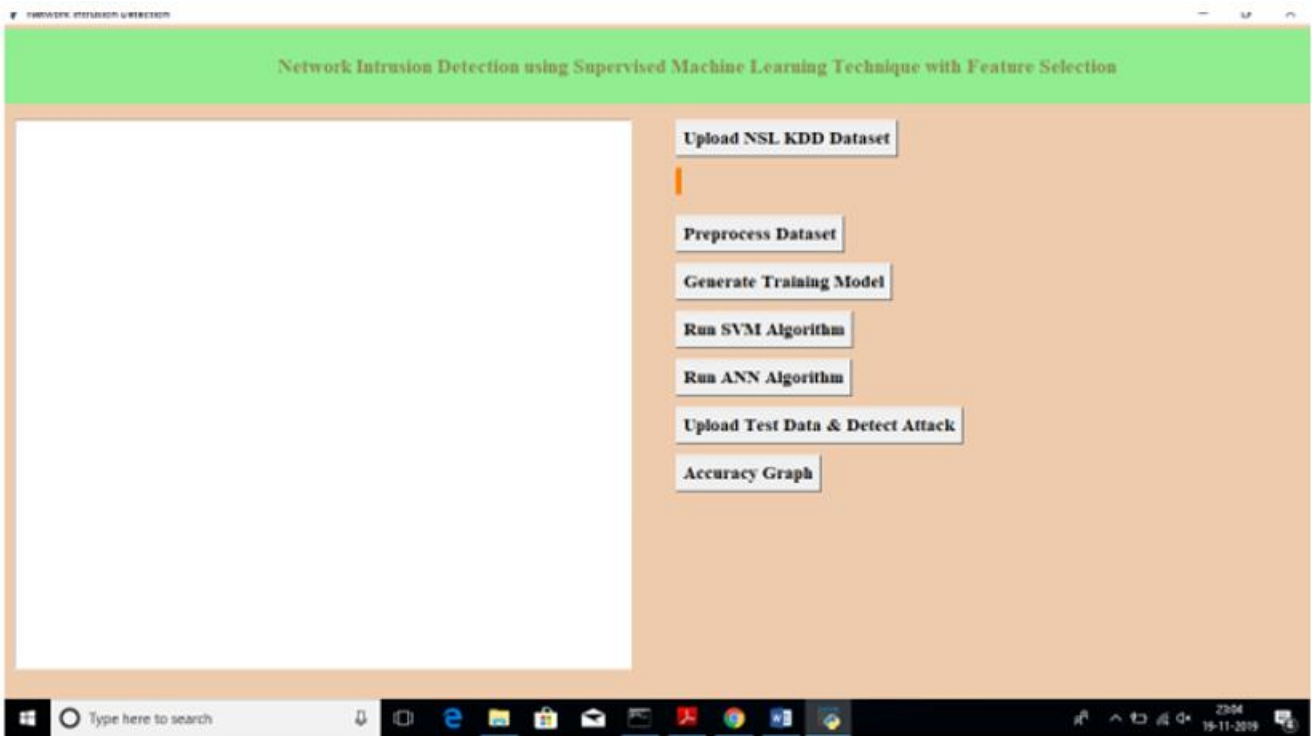
Pass the result through a sigmoid formula to calculate the neuron's output. The Sigmoid function is used to normalize the result between 0 and 1:

$$1/1 + e^{-y}$$

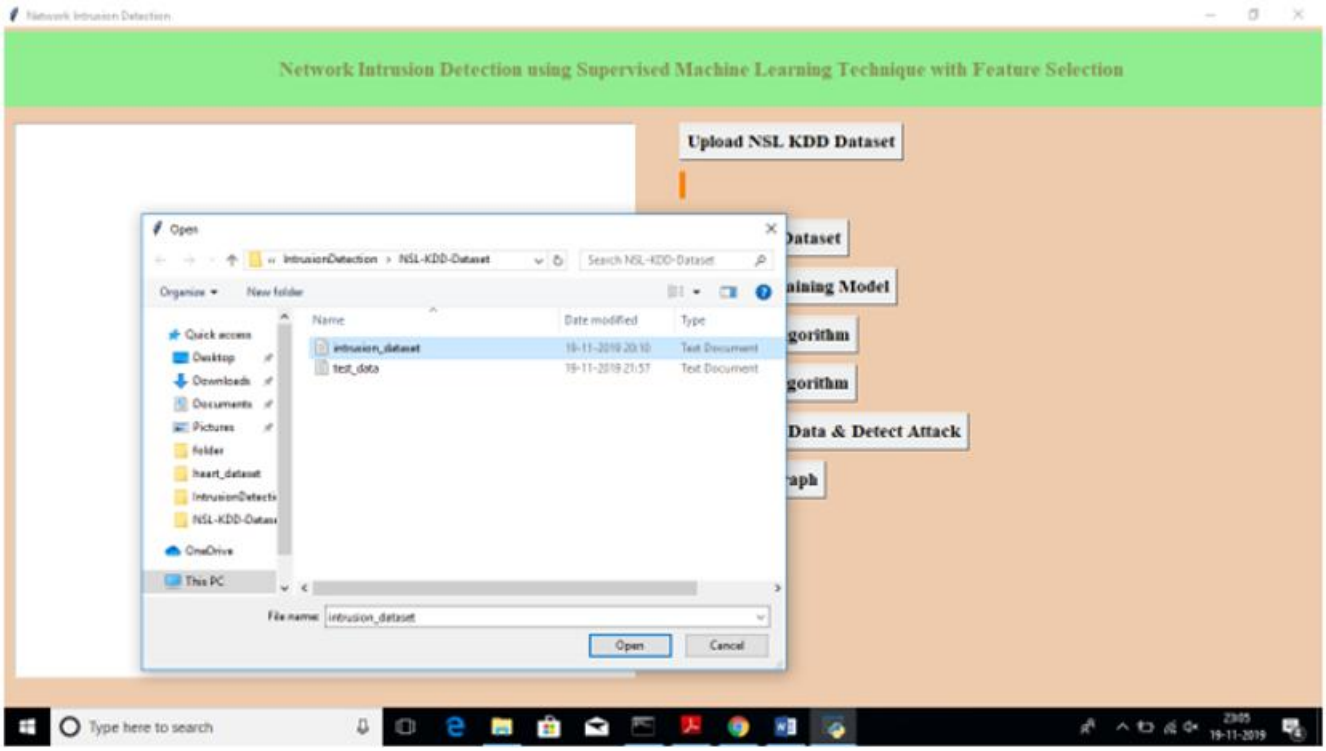
2) Back Propagation
Calculate the error i.e. the difference between the actual output and the expected output. Depending on the error, adjust the weights by multiplying the error with the input and again with the gradient of the Sigmoid curve:

Weight += Error Input Output (1-Output) ,here Output (1-Output) is derivative of sigmoid curve.

Home page:

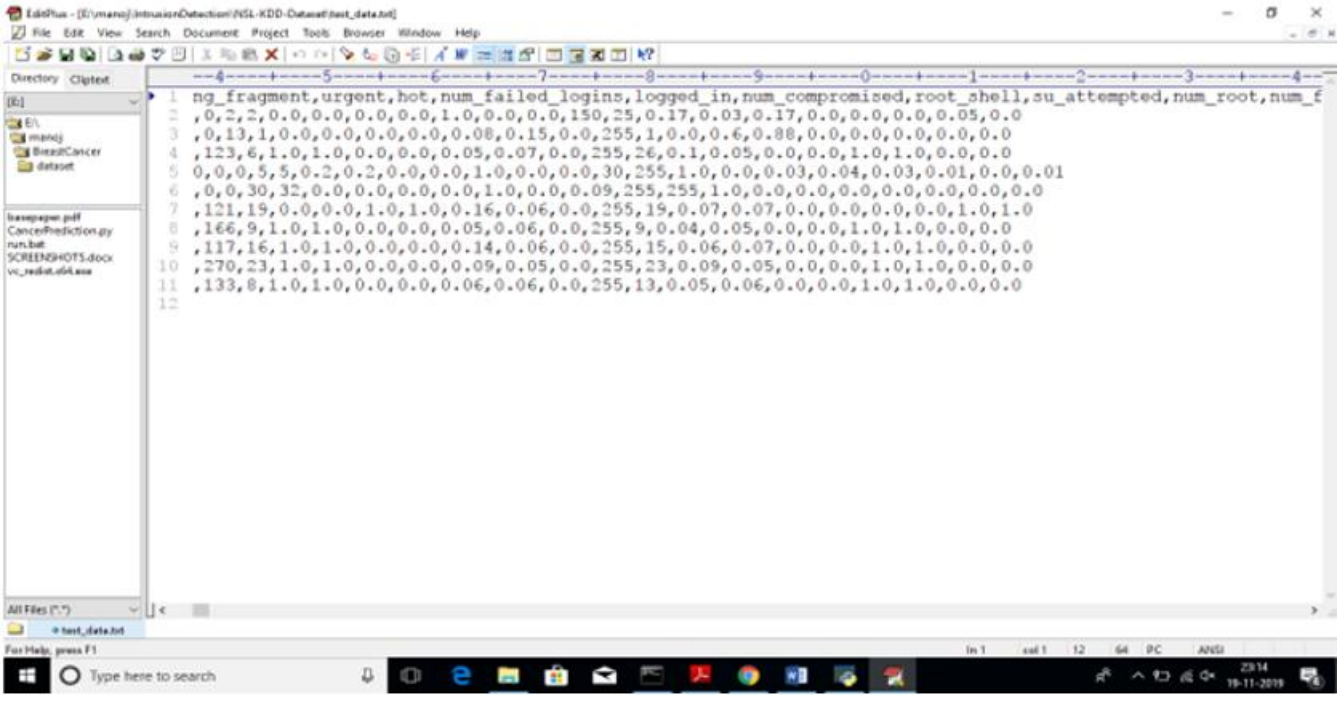


In above screen click on 'Upload NSL KDD Dataset' button and upload dataset.

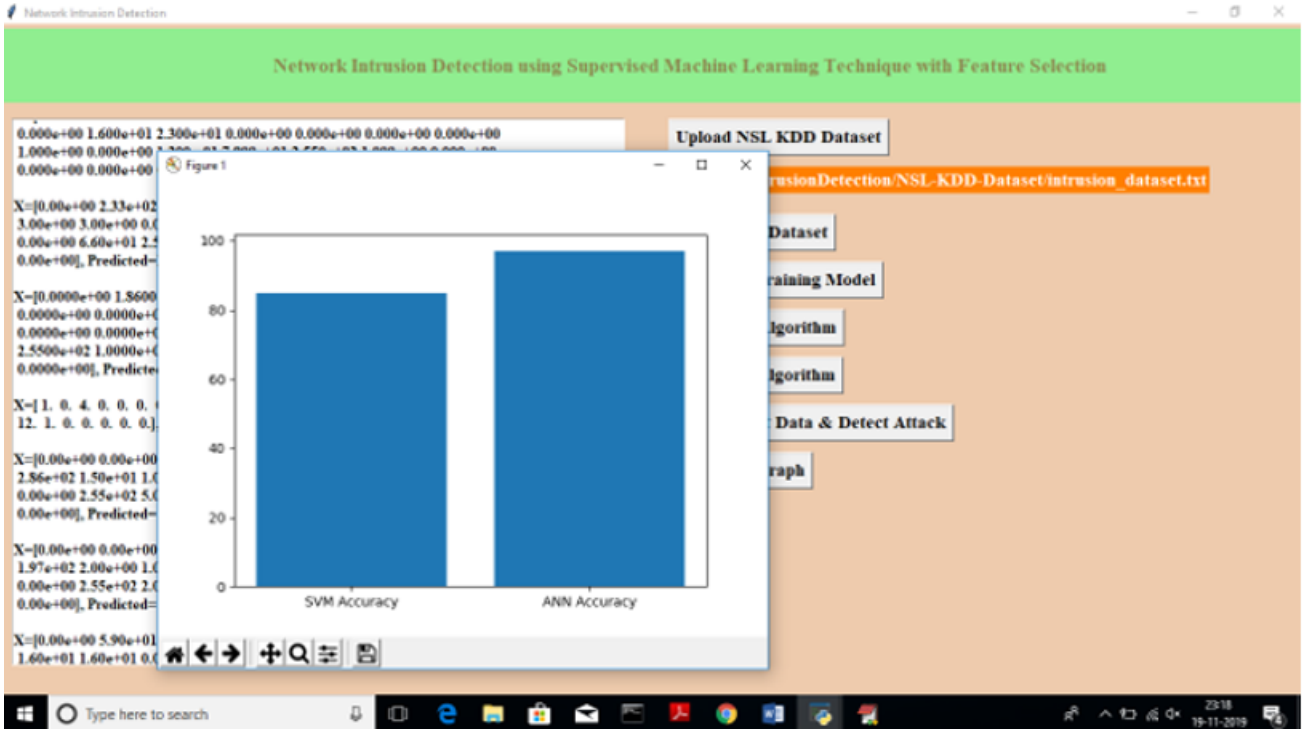


In above screen we got 96.88% accuracy, now we will click on 'Upload Test Data & Detect Attack' button to upload test data and to predict whether test data is normal or contains attack.

All test data has no class either 0 or 1 and application will predict and give us result. See below some records from test data.



In above test data we don't have either '0' or '1' and application will detect and give us result.



From above graph we can see ANN got better accuracy compare to SVM, in above graph x-axis contains algorithm name and y-axis represents accuracy of that algorithms.

5. Discussions

Network Intrusion Detection Systems play a pivotal role in defending against cyber threats like:

- Denial-of-Service (DoS) attacks
- Malware infections
- Unauthorized system access

Unlike firewalls, which operate on predefined rules, NIDS examines deeper traffic patterns and alerts administrators of anomalies.

Detection Approaches:

- **Signature-Based:** Effective for known threats, but blind to new ones
- **Anomaly-Based:** Can detect novel threats, though prone to false positives
- **Hybrid Models:** Combine both methods for enhanced accuracy

6. Conclusion

This study compared supervised learning algorithms for intrusion detection. Key takeaways:

- ANN with feature selection achieved the highest detection accuracy (94.02%)
- Machine learning techniques are powerful tools in identifying both existing and emerging threats
- Feature selection is critical in enhancing model performance

7. Future Scope

Looking ahead, IDS must evolve to handle zero-day threats and reduce false alarms. Future improvements may involve:

- Incorporating AI-driven real-time analysis
- Increasing adaptability to new attack patterns
- Enhancing system responsiveness

Final Thoughts

As cyberattacks become more frequent and complex, deploying intelligent Network Intrusion Detection Systems is no longer optional. While challenges like encrypted traffic and false positives remain, AI and machine learning provide a path toward smarter, more effective cybersecurity.

Author Profile

Mr. Brugumalla Mahendra Achari is completing his MCA at Sri Venkateswara University and has shown a strong interest in Artificial Intelligence. His postgraduate project involved creating a voice-activated assistant (JARVIS) and led to this research initiative under the guidance of: Dr. MOORAMREDDY SREEDEVI.

Dr. Mooramreddy Sreedevi has been serving as an Assistant Professor in the Department of Computer Science at Sri Venkateswara University, Tirupati, Andhra Pradesh, since 2007. She earned her Ph.D. in Computer Science from Sri Venkateswara University, Tirupati. Over the years, she has actively contributed to the academic community. She served as both an Executive Committee Member and the Lady Representative for two years in the SV University Teachers Association. Additionally, she took on the role of UG Examination Coordinator in the university's examination section. Dr. Sreedevi has received three awards for her academic achievements and is the author of nine books. She has published 80 research papers in UGC-reputed journals, participated in 35 international conferences, and attended 54 national conferences. She has also served as a resource person for various universities across India.

References

[1] Song, H., Lynch, M. J., & Cochran, J. K. (2016). *Cyber-victimization analysis.* American Journal of Criminal Justice.
[2] Alaei, P., & Noorbehbahani, F. (2017). *Incremental anomaly-based IDS using limited labeled data.* ICWR 2017.
[3] Saber, M. et al. (2015). *Evaluation methods for IDS.*
[4] Tavallaee, M. et al. (2010). *Credible evaluation of anomaly-based IDS.* IEEE Trans.
[5] Ashoor, A. S., & Gore, S. (2011). *Importance of IDS.*
[6] Zamani, M. & Movahedi, M. (2013). *Machine learning techniques for IDS.* arXiv.
[7] Chakraborty, N. *Comparative study of IDS and IPS.*