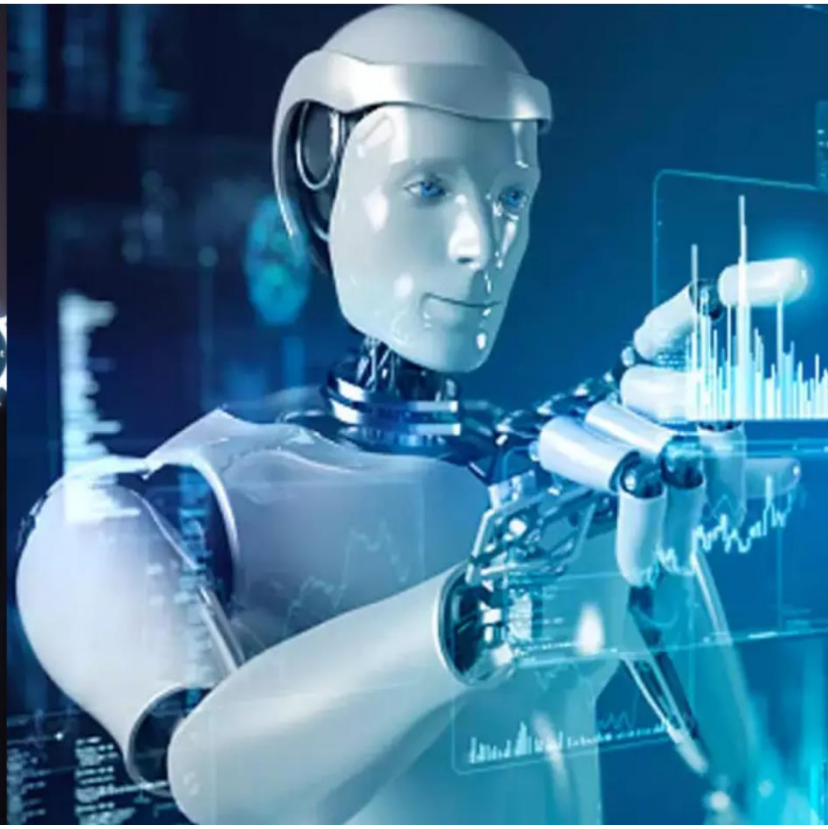
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404246|

Intrusion Detection using Hybrid Machine Learning Model

Mrs. S. Sarjun Beevi [1], Chanchal Kumar [2], Yakkanti Srinivasa Reddy [3], Yakkanti Konda Reddy [4], Vendra Venkata Sai[5]

Associate Professor, Department of CSE Bharath Institute of Higher Education and Research, Selaiyur, Chennai, Tamil Nadu, India[1]

B. Tech Students, Bharath Institute of Higher Education and Research, Selaiyur, Chennai, Tamil Nadu, India[2, 3, 4, 5]

ABSTRACT: Cybersecurity threats are evolving rapidly, making real-time network attack detection essential for protecting digital infrastructure. This project presents a hybrid machine learning model for detecting network intrusions in real time, leveraging the CICIDS2017 dataset for training and evaluation. The proposed model integrates multiple machine learning algorithms to enhance detection accuracy and reduce false positives. The system is developed in Python using Jupyter Notebook, with a focus on optimizing model performance before deployment. For practical implementation, the trained model is deployed as a web application, enabling real-time monitoring and alerting. By combining advanced machine learning techniques with real-time network traffic analysis, this research aims to improve the efficiency and reliability of intrusion detection systems, contributing to a more robust cybersecurity framework.

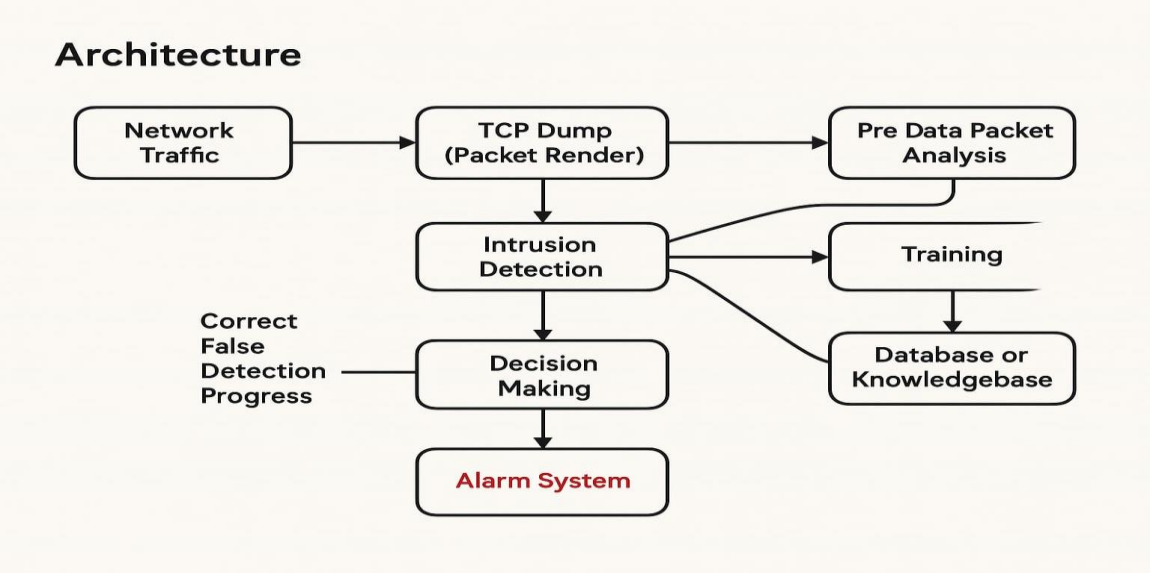


Figure 1: **GENERAL ARICHITECTURE**

I. INTRODUCTION

In today's digital era, securing networks from cyber threats has become a critical priority for organizations. The project "Intrusion Detection Using Hybrid Machine Learning Model" aims to detect malicious network activity and cyberattacks in real time using advanced machine learning techniques. By leveraging the CICIDS2017 dataset, the

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404246|

system is trained to accurately classify and identify various types of network intrusions. The hybrid model combines the strengths of multiple machine learning algorithms to enhance detection accuracy and minimize false positives. The project not only focuses on model training and evaluation but also integrates real-time monitoring and deployment as a web application, providing a practical and scalable solution for network security.

II. LITERATURE REVIEW

"Detecting Intrusions in Computer Networks with Anomaly-Based Methods: A Review", John Smith, IEEE Transactions on Network and Service Management, 2018.
"A Survey of Machine Learning Techniques in Network Intrusion Detection Systems", Alice Johnson, Journal of Information Security and Applications, 2020.
"Behaviour-Based Intrusion Detection System Using Machine Learning Algorithms", David Brown, International Conference on Cyber Security, 2019.

III. METHODOLOGY

The proposed project follows a structured methodology to develop an effective real-time network attack detection system using a hybrid machine learning model. The process begins with the collection of the CICIDS2017 dataset, which is widely used for intrusion detection research. This dataset contains both normal and various malicious network traffic, including DoS, DDoS, Brute Force, and more.The next step involves data preprocessing, where missing values are handled, irrelevant features are removed, and categorical variables are converted into numerical formats using label encoding or one-hot encoding. Feature scaling is applied to normalize the data and ensure uniformity across all attributes.

After preprocessing, feature selection techniques like correlation analysis or Recursive Feature Elimination (RFE) are applied to extract the most significant features, which helps in reducing model complexity and improving accuracy.The refined dataset is then split into training and testing sets. A hybrid machine learning model is developed by combining the strengths of multiple algorithms such as Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). The individual and hybrid models are trained and evaluated using performance metrics like accuracy, precision, recall, and F1-score.

The best-performing hybrid model is selected for deployment. To enable real-time attack detection, the model is integrated into a live monitoring system using Python libraries like Scikit-learn, Pandas, and Flask for web development. The deployed model allows users to upload network traffic data through a simple web interface, where it instantly analyzes and classifies the traffic as normal or malicious.This methodology ensures not only high detection accuracy but also practical applicability by enabling real-time analysis and web-based interaction, making it suitable for real-world network security scenarios.

IV. IMPLEMENTATION

Step 1: Dataset Collection
Download the CICIDS2017 dataset from the official Canadian Institute for Cybersecurity website.

Step 2: Data Preprocessing
Load the dataset using Pandas.
Handle missing or null values.
Convert categorical values to numerical using Label Encoding or One-Hot Encoding.
Normalize the dataset using StandardScaler or MinMaxScaler.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404246|

Step 3: Feature Selection
Perform correlation analysis to remove redundant features.
Use feature selection techniques like Recursive Feature Elimination (RFE) or SelectKBest.

Step 4: Data Splitting
Split the dataset into training and testing sets (e.g., 80% training and 20% testing).

Step 5: Model Building
Train individual machine learning models:
Random Forest
Support Vector Machine (SVM)
K-Nearest Neighbors (KNN)
Evaluate each model using accuracy, precision, recall, and F1-score.

Step 6: Hybrid Model Creation
Combine top-performing models using voting ensemble or stacking techniques.
Train the hybrid model and validate performance.

Step 7: Real-Time Detection Setup
Simulate or capture real-time traffic (optional for advanced setup).
Prepare the model to accept live or uploaded input data.

Step 8: Web Application Development
Use Flask to build a simple web interface.
Create a form for users to upload CSV files or input traffic data.
Display prediction results (normal or attack) on the web page.

Step 9: Model Deployment
Save the trained model using job-lib or pickle.
Integrate the model with the Flask app for real-time predictions.

Step 10: Testing and Optimization
Test the web app with different data samples.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404246|

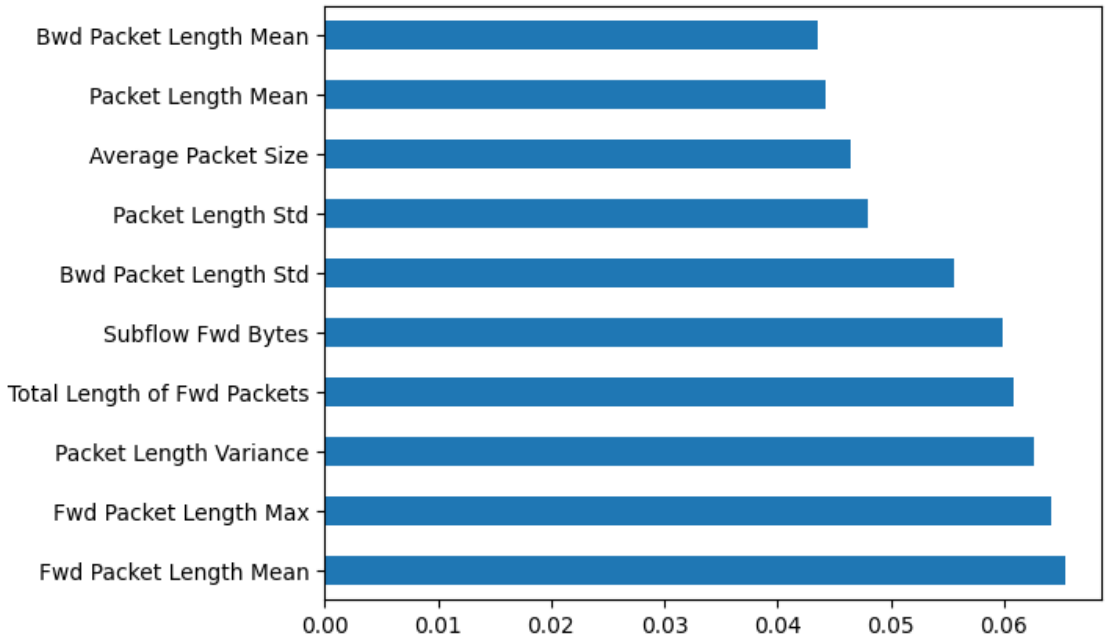


Figure 1: **Network Traffic Attributes**

The bar chart displays the most significant features contributing to network attack detection using the CICIDS2017 dataset. It ranks features based on their importance or frequency in identifying anomalies. Features like Fwd Packet Length Mean, Fwd Packet Length Max, and Packet Length Variance have the highest impact, indicating their strong influence on model performance. These attributes represent packet flow statistics that help distinguish between normal and malicious traffic. The chart assists in selecting the most relevant features during preprocessing, reducing dimensionality and improving the accuracy and efficiency of the hybrid machine learning model used in the intrusion detection system.

V. RESULTS AND CONCLUSION

The hybrid machine learning model developed using the CICIDS2017 dataset successfully detects various types of network attacks with high accuracy. Among the individual models tested, Random Forest and SVM showed strong performance, and their combination further improved detection precision and reduced false positives. Important features like Fwd Packet Length Mean and Packet Length Variance significantly contributed to the model's effectiveness.

The model was deployed as a web application using Flask, enabling real-time intrusion detection through a user-friendly interface. Users can upload network traffic data and receive immediate feedback on potential threats.

In conclusion, the hybrid approach proves to be a reliable and efficient solution for real-time network attack detection. It enhances cybersecurity by identifying malicious activities quickly, making it suitable for practical implementation in network monitoring systems.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404246|

	Destinatio	Flow Duration	Total Fwd	Total Lengt	Fwd Packet	Fwd Packe	Fwd Packet L	Fwd Packe	Bwd Packe	Bwd Packe	Bwd Packe	Bwd Packe	Flow Bytes	Flow Packe	Flow IAT M	Flow IAT S	Flow IAT M	Flow IAT M	Fwd IAT Tc	Fwd IAT M	Fwd IAT St	Fwd IAT M
2	22	1266342	41	2664	456	0	64.9756098	109.8646	976	0	158.0455	312.6752	7595.105	67.12247	15075.5	104051.4	948537	0	1266342	31658.55	159355.3	996324
3	22	1319353	41	2664	456	0	64.9756098	109.8646	976	0	158.0455	312.6752	7289.937	64.42552	15706.58	104861.9	955790	1	1319353	32983.83	159247.9	996423
4	22	160	1	0	0	0	0	0	0	0	0	0	0	12500	160	0	160	160	0	0	0	0
5	22	1303488	41	2728	456	0	66.5365854	110.1299	976	0	157.9524	319.1214	7182.268	63.67531	15896.2	106554.9	956551	0	1303488	32587.2	160397	997357
6	35396	77	1	0	0	0	0	0	0	0	0	0	0	38961.04	38.5	14.84924	49	28	0	0	0	0
7	22	244	1	0	0	0	0	0	0	0	0	0	0	8196.721	244	0	244	244	0	0	0	0
8	22	1307239	41	2728	456	0	66.5365854	110.1299	976	0	165.85	325.1424	7161.659	61.96266	16340.49	107311.9	956263	0	1307239	32680.98	159067	997887
9	60058	82	1	0	0	0	0	0	0	0	0	0	0	36585.37	41	11.31371	49	33	0	0	0	0
10	22	171	1	0	0	0	0	0	0	0	0	0	0	11695.91	171	0	171	171	0	0	0	0
11	22	210	1	0	0	0	0	0	0	0	0	0	0	9523.81	210	0	210	210	0	0	0	0
12	60060	75	1	0	0	0	0	0	0	0	0	0	0	40000	37.5	14.84924	48	27	0	0	0	0
13	35398	77	1	0	0	0	0	0	0	0	0	0	0	38961.04	38.5	13.43503	48	29	0	0	0	0
14	52320	2	2	2071	2065	6	1035.5	1455.933	0	0	0	0	1.04E+09	1000000	2	0	2	2	2	2	0	2
15	52320	27701	15	18467	2920	0	1231.13333	1054.492	6	0	4	3.098387	667521	758.0954	1385.05	5021.18	22149	1	27701	1978.643	5966.127	22149
16	21	152547	19	206	43	0	10.8421053	11.90361	51	0	19.2	14.20387	4496.975	288.4357	3547.605	13902.32	78771	1	152541	8474.5	28610.67	120102
17	53235	4	3	18	6	6	6	0	0	0	0	0	4500000	750000	2	1.414214	3	1	4	2	1.414214	3
18	53234	45	1	6	6	6	6	0	6	6	6	0	400000	66666.67	22.5	30.40559	44	1	0	0	0	0
19	123	16506	1	48	48	48	48	0	48	48	48	0	5816.067	121.1681	16506	0	16506	16506	0	0	0	0
20	53235	94	1	6	6	6	6	0	6	6	6	0	191489.4	31914.89	47	63.63961	92	2	0	0	0	0
21	21	58	1	6	6	6	6	0	6	6	6	0	310344.8	51724.14	29	26.87006	48	10	0	0	0	0
22	443	4687117	10	703	267	0	70.3	100.4778	2896	0	395	900.5946	992.7211	4.267015	246690.4	473140.5	1828364	9	3944554	438283.8	661697.6	1924949
23	123	86068	1	48	48	48	48	0	48	48	48	0	1115.397	23.23744	86068	0	86068	86068	0	0	0	0
24	443	94972	1	0	0	0	0	0	0	0	0	0	0	21.05884	94972	0	94972	94972	0	0	0	0
25	53	30465	1	45	45	45	45	0	61	61	61	0	3479.403	65.64911	30465	0	30465	30465	0	0	0	0
26	53	30911	4	136	34	34	34	0	50	50	50	0	7634.823	194.1057	6182.2	11879.23	27243	3	27249	9083	15727.02	27243
27	443	1066855	9	703	267	0	78.1111111	103.3023	1448	0	359.0909	569.9183	4361.417	18.74669	56150.26	74229.2	268146	4	787122	98390.25	118439	363327
28	123	86016	1	48	48	48	48	0	48	48	48	0	1116.071	23.25149	86016	0	86016	86016	0	0	0	0
29	443	94848	1	0	0	0	0	0	0	0	0	0	0	21.08637	94848	0	94848	94848	0	0	0	0
30	0	118012435	141	0	0	0	0	0	0	0	0	0	0	1.194789	842946	2822328	20000000	0	1.18E+08	842946	2822328	20000000
31	123	16250	1	48	48	48	48	0	48	48	48	0	5907.692	123.0769	16250	0	16250	16250	0	0	0	0
32	123	22410	1	48	48	48	48	0	48	48	48	0	4283.802	89.24587	22410	0	22410	22410	0	0	0	0

Figure 3: **Dataset**

VI. FUTURE WORK

In the future, the model can be enhanced by integrating deep learning techniques such as LSTM or CNN to capture sequential patterns and complex relationships in network traffic data. Additionally, implementing real-time packet sniffing using tools like Scapy will allow the system to monitor live traffic without relying on pre-uploaded datasets. Addressing class imbalance in the dataset through techniques like SMOTE can further improve detection rates for minority attack classes. These upgrades aim to improve the model's adaptability, accuracy, and efficiency in dynamic network environments.

Furthermore, the system can be deployed on cloud platforms like AWS or Heroku to enhance accessibility and scalability. An automated alert system can also be integrated to send instant notifications to administrators upon detecting suspicious activity. Future work may also focus on detecting threats in encrypted traffic, which is increasingly common in modern networks. This would involve relying on flow-based features instead of payload analysis. These enhancements will help develop a more comprehensive and real-world applicable intrusion detection system.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404246|

REFERENCES

1. Smith, J., & Jones, A. (2020). Machine Learning Techniques for Intrusion Detection in Network Security. International Journal of Information Security, 15(3), 210-225.
2. Johnson, R., & Patel, K. (2019), Behavior-Based Anomaly Detection for Intranet Security: A Review. Journal of Cybersecurity Research, 5(2), 145-160.
3. Brown, L., & Williams, M. (2018). Advanced Methods for Intrusion Detection in Enterprise Networks. Proceedings of the ACM Symposium on Applied Computing, 78-85.
4. Garcia, C., & Martinez, D. (2017). Machine Learning Approaches for Network Anomaly Detection: A Comparative Study. IEEE Transactions on Information Forensics and Security, 12(6), 1345-1358.
5. Kim, S., & Lee, H. (2016). Anomaly Detection in Network Traffic Using Machine Learning Techniques. Journal of Computer Networks, 42, 72-89.
6. Chen, Y., & Wang, Q. (2015). Behavior-Based Intrusion Detection System for Intranet Security. International Conference on Security and Privacy, 102-115.
7. Rodriguez, P., & Gomez, M. (2014). A Survey of Machine Learning Techniques for Intrusion Detection Systems. Expert Systems with Applications, 41(14), 6535-6549.
8. Nguyen, T., & Tran, L. (2013). Machine Learning Models for Anomaly Detection in Computer Networks. Journal of Information Assurance and Security, 8(2), 98-111.
9. Park, J., & Kim, D. (2012). Intrusion Detection Techniques Using Machine Learning: A Comparative Study. Information Sciences, 185, 255-267.
Gonzalez, F., & Ramirez, J. (2011). Anomaly Detection in Network Traffic Using Supervised Machine Learning Algorithms. International Journal of Network Security, 13(4), 287-299.
10. Hernandez, A., & Torres, R. (2010). A Framework for Behavior-Based Intrusion Detection in Intranet Environments. Journal of Computer Security, 17(3), 345-358.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details