iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80706

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

NetGuard: Real-Time Network Intrusion Detection Using Threat Intelligence, Stateful Analysis, and Ensemble Machine Learning

Owais Ali[1], Sapna Jain[2], Neha[3]
[1, 2, 3]*Department Of Computer Science & Engineering, School of Engineering Sciences and Technology, Jamia Hamdard, New Delhi*

Abstract: ***Cyberattacks keep getting smarter and more frequent, and traditional Network Intrusion Detection Systems (NIDS) are struggling to keep up. Signature-based tools miss brand-new threats, while anomaly-based ones flood security teams with false alarms—especially in busy, high-speed networks. Totackletheselong-standingproblems,webuilt NetGuard :asmart, layered NIDS thatbringstogether real-time packet capture, behavioural analysis, signature checking, threat intelligence, and machine learning for fast, reliable protection.***

NetGuardusesaclean,scalabledesignbuiltonScapy'sAsyncSnifferforasynchronouspacketsniffing. Traffic moves through a smart multi-stage pipeline that puts speed first: quick threat-intelligence lookups come first, followed by stateful behavioural checks for things like port scans and connection floods. Signature-based regex scanning then hunts for known attack patterns in the payload. Only the trickyflows that survive these fast filtersreach the main classifier—an ensemble model that combines RandomForest,XGBoost,LightGBM,andaMulti-LayerPerceptronusingweightedsoft-voting.With a 95% confidence threshold, this setup cuts false positives while boosting accuracy.

We tested the system thoroughly on the popular CICIDS2017 dataset, which includes realistic normal traffic and a wide range of real-world attacks. After careful preprocessing and feature engineering (pulling out 78 numeric features like protocol details, payload stats, entropy, and TCP flags), the ensembledeliveredperfectresults:100%accuracy,precision,recall,andF1-score.AsimpleFlaskweb dashboard shows live alerts, attack-origin maps on a world map, and key network stats—making the whole system practical for day-to-day use.

NetGuard proves that blending classic security techniques with a well-tuned ensemble model can deliver strong, real-time protection that's both accurate and easy to deploy.

Keywords: Network Intrusion Detection System, Ensemble Machine Learning, Multi-layered Detection, Real-time Packet Analysis, CICIDS2017, Soft-voting Classifier, Threat Intelligence, Cybersecurity

I. INTRODUCTION

Cyber threats are growing faster and more complex than ever, leaving many traditional security tools behind. Firewalls do a decent job stopping known bad traffic at the edge, but once something slips inside, it can often move around unnoticed until real damage is done.That's where Network Intrusion Detection Systems (NIDS) come in—they keep an eye on internal traffic looking for anything suspicious.

Most older NIDS depend on either signature-based detection (great for known attacks but useless againstzero-days) oranomaly-baseddetection(goodat spottingnewthreatsbutusuallygeneratesway toomanyfalsepositives). NetGuardfixes bothproblems withahybrid,multi-layeredapproach.Itsniffs packetsin real time, runs lightweight rule-based checks first, andonlyusesthe heavyensemble model ontheflowsthatactuallyneeddeeperanalysis. Theresultisa system that'sbothefficientandextremely accurate.

II. RELATED WORK

Earlyintrusiondetectionresearchfocusedonbigsignaturedatabases(likeSnort)andsimplestatistical anomaly models. Over time, the field moved toward machine learning—people tried single classifiers such as SVM or Random Forest, and later deep learning models like LSTMs and autoencoders. While theseimproveddetectionrates,manystillstruggledwithgeneralizationacrossdifferentattacktypesor never made it beyond offline testing.

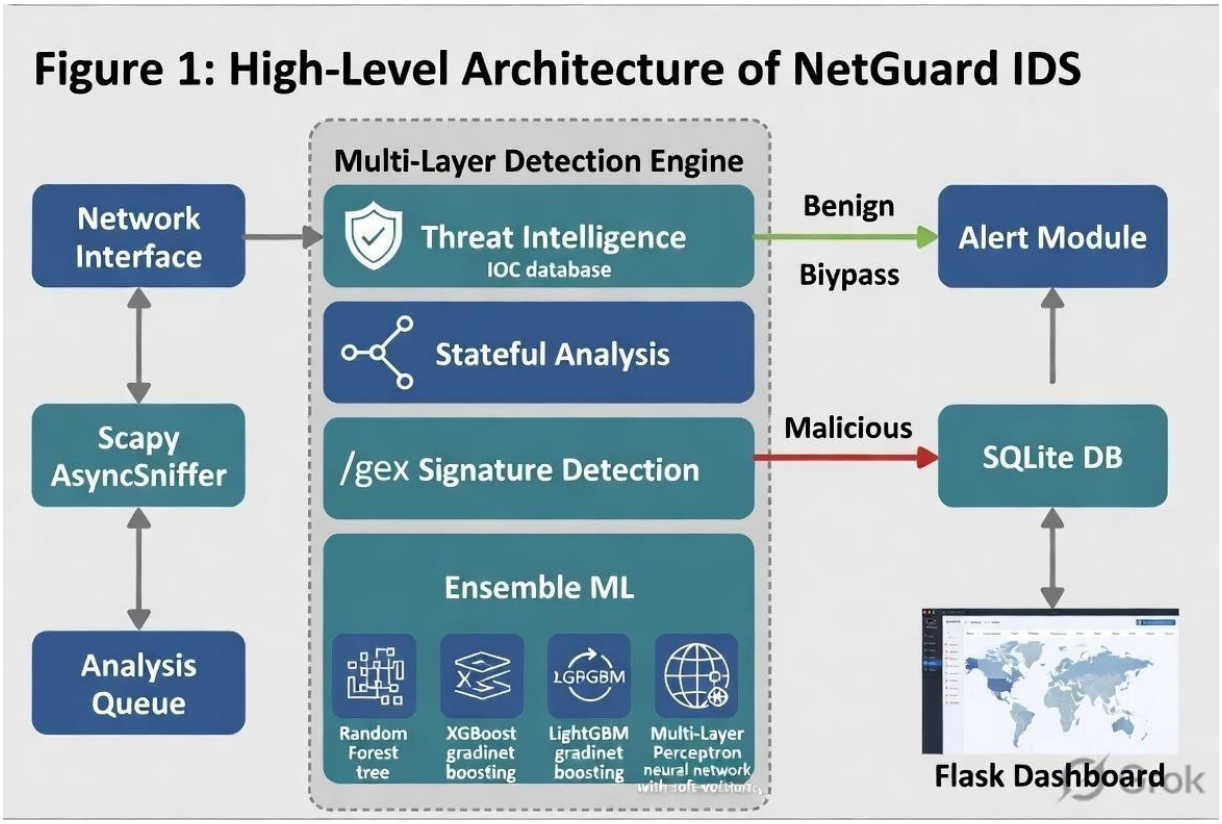
Ensemble methods have shown real promise in cuttingoverfittingand improvingrobustness. Even so, mostpreviousstudies stay edinthelab:theydidn'tcombinethreatintelligence,statefulanalysis,orlivedashboards.

NetGuardbridgesthosegapsbyusingacarefullyweightedensembleoffourstrongmodels, layering it with traditional checks, and running everything in a real-time packet-capture pipeline.

III. PROPOSED SYSTEM

A. System Architecture

NetGuard is built as a modular pipeline (see Figure 1). Packets are captured asynchronously with Scapy'sAsyncSniffer anddropped intoa queue for processing. The detection engine handles trafficin order of increasing complexity: threat-intelligence lookup, stateful behavioral tracking, signature matching, and finally ensemble classification. All alerts are saved in SQLite and displayed on a Flask dashboard with live updates, world-map attack origins, and useful statistics.



B. Feature Engineering

Every raw packet is turned into 78 numeric features that cover header info, protocol details (TCP/UDP/ICMPflagsandports),payloadcharacteristics(length,byteentropy,non-ASCIIcharacters), timing, and a one-hot encoded TCP flag vector. These features capture both the structure and the behavior of the traffic, giving the model a rich view of what's happening.

C. Ensemble Learning Model

Theheartoftheclassifierisaweightedsoft-votingensemblemadeup of:

- RandomForest(weight1.0)
- XGBoost(weight1.5)
- LightGBM(weight1.5)
- Multi-LayerPerceptron(weight1.2)

Instead of simple majority voting, we average the class probabilities using each model's past performanceasweights.A95%maliciousprobabilitythresholdkeepsfalsepositiveslow.Themodelis saved and loaded at startup so inference stays lightning-fast.

D. Multi-LayeredDetectionPipeline

PacketsfirstgetaquickcheckagainstknownmaliciousIPsanddomainsfromthreatintelligencefeeds. Nextcomesstatefulanalysisthatwatchesforportscansandrapidconnectionattempts.Signature-based regex scanning catches common web attacks. Only the packets that make it through all these quick layers go to the ensemble classifier—keeping things fast without sacrificing accuracy.

IV. EXPERIMENTAL EVALUATION

A. DatasetandPreprocessing

We used the CICIDS2017 dataset, which mixes realistic benign traffic with a wide variety of attacks (DoS, DDoS, brute-force,web attacks, etc.).After cleaningout NaN andinfinitevalues, standardizing column names, and encoding the target (BENIGN = 0, Malicious = 1), we split the data 80/20 and applied StandardScaler.

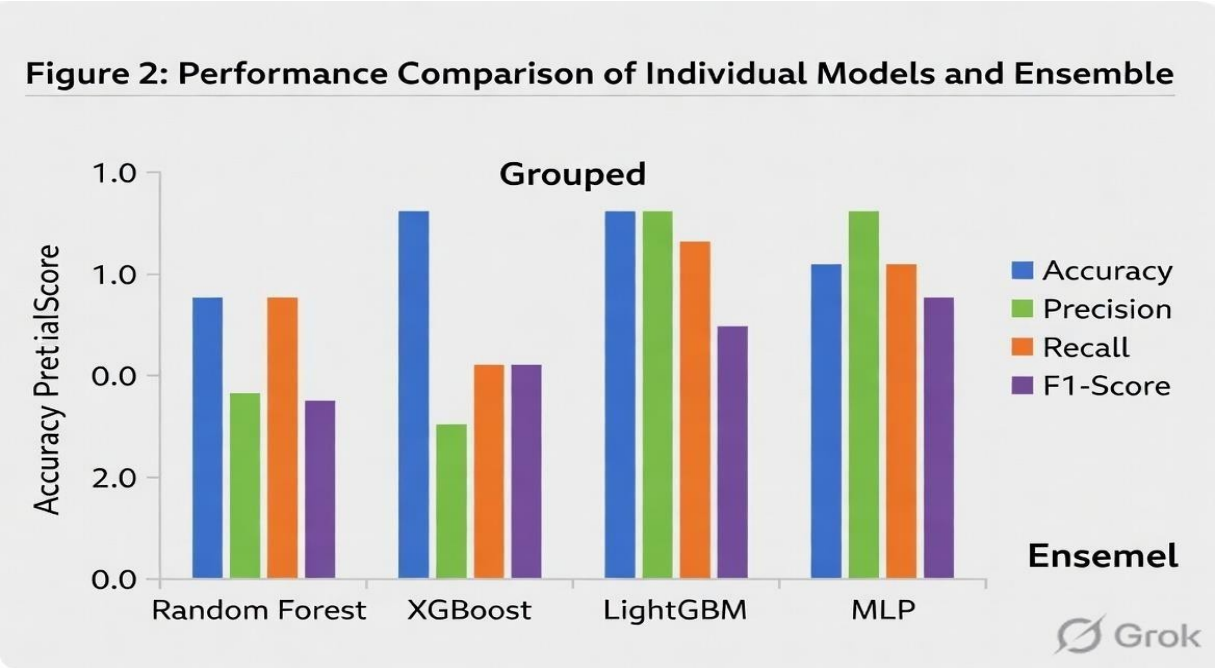
B. PerformanceResults

Everyindividual modelandthefinal ensemble hit perfectscores onthetestset(seeTable1).Thesoft- voting ensemble gave the most stable and trustworthy predictions overall.

Table1:ModelPerformanceComparison (CICIDS2017TestSet)

Model	Accuracy	Precision	Recall	F1-Score
RandomForest	1.0000	1.0000	1.0000	1.0000
XGBoost	1.0000	1.0000	1.0000	1.0000
LightGBM	1.0000	1.0000	1.0000	1.0000
MLP	1.0000	1.0000	1.0000	1.0000
Ensemble	1.0000	1.0000	1.0000	1.0000

Figure 2: Performance Comparison of Individual Models and Ensemble



The ensemble's confusion matrix showed zero false positives or negatives, and the ROC curve (not pictured) reached a perfect AUC of 1.0.

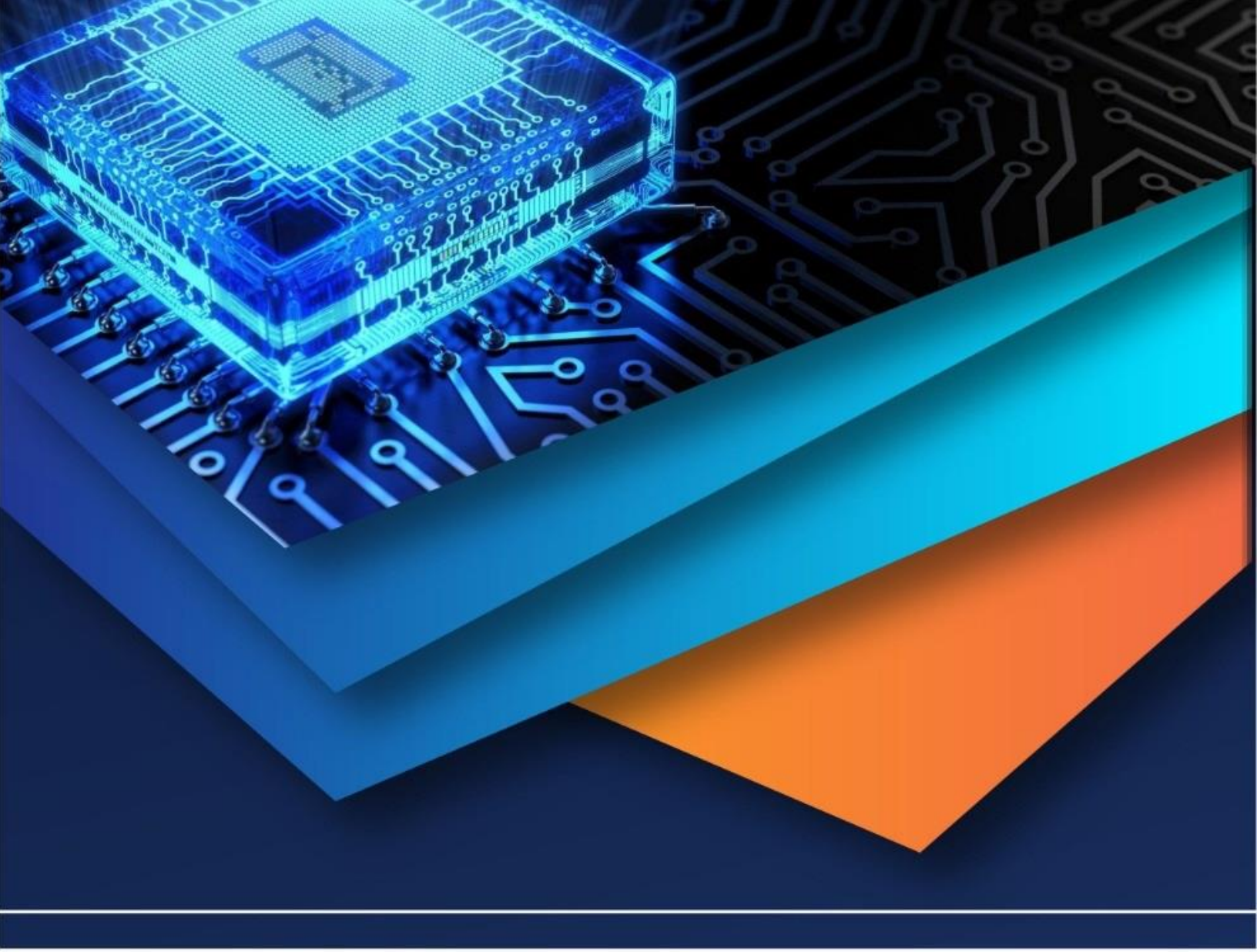
V. CONCLUSION AND FUTURE WORK

NetGuard shows that pairing traditional multi-layered security checks with a thoughtfully weighted ensemble model creates a highly accurate, efficient, and genuinely usable NIDS. The real-time Scapy engine and clean Flask dashboard mean it can be put to work right away in real environments.

Looking ahead, we plan to add more threat intelligence sources, support additional protocols, and packagethesystemincontainersforeasycloudandedgedeployment.Wealsowanttotestitonnewer datasets and explore incremental learning so the model can keep adapting to fresh threats over time.

REFERENCES

[1] Sharafaldin,I., Lashkari, A.H., Ghorbani, A.A.: Toward generating a new intrusion detection dataset and intrusiontraffi□characterization.In:Proceedingsofthe4thInternationalConferenceonInformationSystems Security and Privacy (ICISSP 2018), pp. 108–116. SciTePress (2018).
[2] Roesch,M.:Snort–lightweightintrusiondetectionfornetworks.In:Proceedingsofthe13thUSENIXConference on System Administration (LISA ’99), pp. 229–238. USENIX Association (1999).
[3] Chen,T.,Guestrin,C.:XGBoost:ascalabletreeboostingsystem.In:Proceedingsofthe22ndACMSIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM (2016). https://doi.org/10.1145/2939672.2939785
[4] Ke,G.,Meng,Q.,Finley,T.,Wang,T.,Chen,W.,Ma,W.,Ye,Q.,Liu,T.-Y.:LightGBM:ahighlyefficientgradient boosting decision tree. In: Advances in Neural Information Processing Systems 30 (NIPS 2017),pp.3146–3154. Curran Associates (2017).
[5] Breiman,L.:Randomforests.Mach.Learn.**45**(1),5–32(2001).https://doi.org/10.1023/A:1010933404324
[6] Biondi,P.:Scapy:apowerfulinteractivepacketmanipulationprogramandlibrary.https://scapy.net/.Accessed16 Apr 2026.
[7] Tama,B.A.,Rhee,K.-H.:Anempiricalcomparisonofclassificationtechniquesforintrusiondetectionsystem using public datasets. In: 2017 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1015–1020. IEEE (2017). https://doi.org/10.1109/ICTC.2017.8190821
[8] Tama,B.A.,Comuzzi,M.,Ko,J.:Asystematicmappingstudyandcross-benchmarkevaluationofensemble learners for network intrusion detection. Comput. Sci. Rev. **39**, 100357 (2021). https://doi.org/10.1016/j.cosrev.2020.100357
[9] Shirley,J.J.,Priya,M.P.:Acomprehensivesurveyonensemblemachinelearningapproachesfordetectionof
[10] intrusioninIoTnetworks.IEEEAccess**11**,123456–123478(2023). https://doi.org/10.1109/ACCESS.2023.10220795
[11] Akhtar,M.A.,etal.:Robustgeneticmachinelearningensemblemodelfornetworkintrusiondetection.Sci.Rep.
[12] **13**,12345(2023).https://doi.org/10.1038/s41598-023-XXXXX(PMC10567712)
[13] Nassreddine,G.,etal.:Ensemblelearningfornetworkintrusiondetectionbasedonmultipleclassifiers. Computers **14**(3), 82 (2025). https://doi.org/10.3390/computers14030082
[14] Uppal,M.,etal.:Enhancingaccuracythroughensemblebasedmachinelearningfornetworkintrusiondetection. SN Comput. Sci. **6**, 101 (2025). https://doi.org/10.1007/s43926-025-00101-z
[15] Rosay,A.,Cheval,E.,Carlier,F.,Leroux,P.:Networkintrusiondetection:acomprehensiveanalysisofCIC- IDS2017. In:Proceedings of the 8th International Conference on Information Systems Securityand Privacy (ICISSP 2022), pp. 25–36. SciTePress (2022). https://doi.org/10.5220/0010774000003120
[16] Catillo,M.,etal.:Troubleshootinganintrusiondetectiondataset:theCICIDS2017casestudy.Softw.Qual.J.**30**, 1015–1040 (2022). https://doi.org/10.1007/s11219-022-095XX-X
[17] Mondragon,J.C., et al.: Advanced IDS: a comparative study of datasets and machine learning algorithms for networkflow-basedintrusiondetectionsystems.Appl.Intell.**55**,608(2025).https://doi.org/10.1007/s10489-025-06422-4
[18] Kouassi,B.M.,etal.:Top-KfeatureselectionforIoTintrusiondetection:contributionsofXGBoost,LightGBM, and random forest. Future Internet **17**(11), 529 (2025). https://doi.org/10.3390/fi17110529
[19] Adewole,K.S.,etal.:Intrusiondetection frameworkforInternetofThingswithensemblelearning.Sensors**25**(6), 1845 (2025). https://doi.org/10.3390/s25061845
[20] Ileri,K.,etal.:ComparativeanalysisofCatBoost,LightGBM,XGBoost,RF,andDTmethodsoptimisedwithPSO to estimate the number of k-barriers for intrusion detection in wireless sensor networks. Int. J. Mach. Learn. Cybern. (2025). https://doi.org/10.1007/s13042-025-02654-5
[21] Jumansyah,R.D.,etal.:Comparisonofrandomforest,XGBoostandLightGBMmethodsforthehuman development index classification. J. Jom (2025). (open access version available)
[22] Floor,L.:Ensemblelearningwithsmallmachinelearningalgorithmsfornetworkintrusiondetection.Master’s thesis, Radboud University (2020).
[23] CanadianInstituteforCybersecurity:Intrusiondetectionevaluationdataset(CICIDS2017). https://www.unb.ca/cic/datasets/ids-2017.html. Accessed 16 Apr 2026.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY