



IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

www.ijitce.com



Email : ijitce.editor@gmail.com or editor@ijitce.com

ENSEMBLE-LEARNING-BASED DEEP NEURAL NETWORK ATTACK CLASSIFICATION OF IMBALANCED IOT INTRUSION DATA

AMBATI VAISHNAVI¹, B RAMA GANESH², A DHANASEKHAR REDDY³, K LAKSHMAN KUMAR⁴

¹P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,
Email: vaishnavivr1207@gmail.com

²Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email:
ramaganesh34@gmail.com

³Assistant Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology,
Puttur, Email: ghanasekhar918@gmail.com

⁴Assistant Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur,
Email: lakshman5804@gmail.com

ABSTRACT: IoT gadgets feature the requirement for solid safety efforts to lessen weaknesses and risks in connected networks. This paper presents a Bagging Classifier (BC)-based Deep Neural Network (DNN) strategy to deal with class irregularity worries in IoT intrusion detection datasets. This strategy utilizes deep learning and ensemble learning to improve intrusion detection and arrangement. Four unmistakable ID datasets — NSL-KDD, KDDCUP99, UNSW-NB15, and Bot-Io — show promising accuracy, precision, recall, F-score, and false positive rate. The recommended technique beats past strategies, particularly while involving 10 base assessors in the bagging ensemble approach. Further exploration utilizes Convolutional Neural Networks (CNN) and hybrid CNN + Long Short-Term Memory (LSTM) models to accomplish close to 100% accuracy. Flask is utilized to give a front-end connection point to user testing and verification, further developing convenience. This exploration further develops IoT ID by showing ensemble learning's capacity to deal with class unevenness issues and further develop network security.

Index Terms—Bagging, class imbalance, class weights, deep neural network (DNN), ensemble

learning, Internet of Things (IoT), intrusion detection system (IDS).[33]

1. INTRODUCTION:

Internet of Things (IoT) networks currently have unmatched data access and conduct unconventionality [1], [2]. The development of organization and online applications has expanded information volume and organization weaknesses and dangers. The quick organization of IoT gadgets and the natural plan of IoT networks have convoluted the security climate, bringing a few challenges and risks [1], [3].

Huge endeavors have been made to make successful IoT Intrusion Detection Systems (IDS) to address these hardships. These IDS look at and classify network information tests into assault and customary traffic [3]. Deep Learning (DL) calculations are promising for ID and classification systems because of their mind complex learning abilities and far and wide use across application spaces [4].

Intrusion detection systems datasets for IoT networks help innovative work. UNSW-NB15 and Bot-IoT are models [5]. These measurements

practically portray network traffic and incorporate a few IoT danger regions.

Interruption discovery frameworks on IoT networks battle with dataset class lopsidedness. In the BoT-IoT dataset, only 0.013% of information examples comprise normal organization traffic [6]. These lopsided class disseminations could misshape order calculations and corrupt IDS execution in IoT organizations.

These troubles roused our examination to defeat class unevenness in IoT ID datasets. The Bagging Classifier (BC)- based Deep Neural Network is our ensemble learning technique. This strategy further develops IoT ID and order utilizing DL and ensemble learning.

This review tests the BC-based DNN procedure for class awkwardness in IoT ID datasets. The proposed method will be tried utilizing four unmistakable ID datasets: NSL-KDD, KDDCUP99, UNSW-NB15, and BoT-IoT.

The paper's technique, exploratory arrangement, discoveries, and analysis follow this presentation. We will likewise look at our strategy other class irregularity ways to deal with show its convenience in further developing IoT ID and classification.

In the accompanying parts, we will cover the important work, the procedure utilized in this exploration, the exploratory outcomes, and the ends and future examination.[35]

2. LITERATURE SURVEY

Lately, research on deep learning (DL) in intrusion detection systems (IDS) for network security has developed. This part audits critical exploration in

this theme, featuring significant commitments and discoveries.

DL in IDS was widely covered by Aminanto and Kim (2016) [1]. Their review examined DL models and ID techniques, featuring their advantages and disadvantages. This work laid out the establishments for future examination on DL's capacity to further develop IDS.

Thakkar and Lohiya (2020) looked at feature selection attack arrangement [2]. They tried different feature selection techniques to further develop IDS accuracy and proficiency. The work upgraded feature representation for ID by evaluating feature selection techniques.

Likewise, Thakkar and Lohiya (2020) analyzed swarm and developmental calculations in IDS [3]. Their examination demonstrated the way that multitude and developmental calculations can further develop IDS execution, quite in adaptability and variation to dynamic organization settings. This exploration uncovered better approaches to fortify IDS against arising dangers.

Thakkar and Lohiya (2021) analyzed DL-based IDS regularization combination [4]. Their examination analyzed how regularization techniques further develop ID DL model speculation and versatility. The review decreased overfitting and further develop IDS unwavering quality by evaluating regularization techniques.

Lohiya and Thakkar (2021) likewise explored IoT application regions, evaluation datasets, and research issues [5]. Their broad examination showed that IoT networks present one of a kind security issues and focused on the requirement for viable ID arrangements intended for them. This study showed

that IoT-based risks require explicit datasets and assessment techniques.

IoT network scientific investigation requires sensible botnet datasets, as indicated by Koroniotis et al. (2019) [6]. The Bot-IoT dataset, which reproduces botnet movement in IoT networks, was their primary undertaking. The review further developed IoT network security and scientific examination by organizing a specific dataset.

In their weighty Deep Learning book, Goodfellow et al. (2016) covered DL techniques [7]. This legitimate aide makes sense of DL ideas and applications in network security and ID. The book has molded DL-based IDS research by making sense of key standards and techniques.

Likewise, Dong and Wang (2016) thought about DL and standard network intrusion detection advances [8]. Their examination analyzed DL approaches' accuracy, proficiency, and versatility to customary techniques. The review uncovered the advantages and disadvantages of DL in ID by testing DL models against standard techniques.

The writing survey shows that IDS have gained significant headway utilizing DL draws near. Scientists have progressed ID and network security through exhaustive outlines, relative examinations, and dataset age. These examinations give suggestions and techniques for making fruitful IDS for current network threats.

3. METHODOLOGY

a) Proposed work:

The recommended concentrate on utilizes a Bagging Classifier (BC)- based Deep Neural Network (DNN) for IoT intrusion detection and classification. This

technique utilizes DL and ensemble learning out how to address ID dataset class unevenness. Ten DNN models are utilized as premise assessors for bagging, with class loads to adjust class appropriation. The task likewise utilizes a CNN and a CNN+LSTM model, accomplishing remarkable accuracy. The framework's frontend utilizes the CNN+LSTM model, which has almost 100% accuracy with KDDCUP99. The Flask framework is utilized to give an easy to understand connection point to client testing and collaboration, keeping up with safe access through client confirmation and further developing Intrusion Detection System security.[37]

b) System Architecture:

Data discovery and investigation to comprehend intrusion detection datasets start the system architecture. Data visualization grasps information dissemination and relationships. Clean, preprocess, and change the information for examination utilizing information handling techniques. ID models use feature selection ways to deal with track down important attributes.

The intrusion detection architecture utilizes DNN and LSTM models. These models utilize indicated qualities to sort network traffic as typical or assault. Model execution in recognizing different dangers is assessed utilizing execution measures.

Real-time attack detection and moderation are additionally included. To give a dependable and effective Intrusion Detection System (IDS) for IoT networks, the plan incorporates information driven techniques, ML models, and execution evaluation strategies.

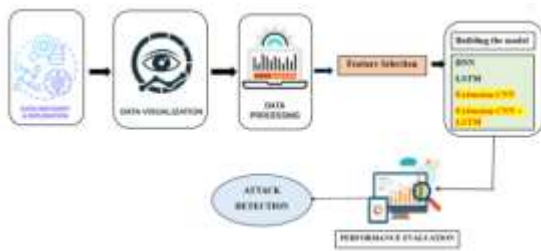


Fig 1 Proposed Architecture

c) Dataset collection:

Four particular intrusion detection datasets intended for IoT networks— KDDCUP99, NSL KDD, UNSW-NB15, and BoT-IoT — make up the undertaking's information assortment. The 1999 DARPA ID Assessment Program yielded the KDDCUP99 dataset, which fills in as a norm for surveying interruption location frameworks under fluctuating assault situations and with regular network traffic.

id	dur	proto	service	state	spkts	dsts	dstbytes	dstbytes	rate	...	cl_dst_sport_bin	cl_dst_svc_bin	is_top_k
1	0.000011	udp	-	INT	2	0	486	0	90900.0002	...	1	2	
2	0.000008	udp	-	INT	2	0	1762	0	125000.0000	...	1	2	
3	0.000005	udp	-	INT	2	0	1988	0	200000.0059	...	1	3	
4	0.000006	udp	-	INT	2	0	900	0	180000.0008	...	1	3	
5	0.000010	udp	-	INT	2	0	2126	0	100000.0025	...	1	3	

Fig 2 data set

NSL KDD is a superior rendition of the KDDCUP99 dataset that might be utilized for training and testing IDS models since it has less overt repetitiveness and better portrayal. UNSW-NB15 works with the improvement of intrusion detection systems tweaked for IoT settings by offering a complete assortment of genuine organization traffic information, including attacks exceptional to the Internet of Things.[39]

duration	protocol_type	service	flag	wrc_bytes	dst_bytes	land	wrong_fragment	urgent	hot	...	cl_dst_svc_bin	cl_dst_sport_bin
0	0	tcp	http	SF	181	8450	0	0	0	...	0	0
1	0	tcp	http	SF	226	480	0	0	0	...	0	0
2	0	tcp	http	SF	236	1017	0	0	0	...	0	0
3	0	tcp	http	SF	219	1337	0	0	0	...	0	0
4	0	tcp	http	SF	207	2032	0	0	0	...	0	0

Fig 3 data set

At long last, the BoT-IoT dataset offers a practical and requesting dataset for evaluating IDS viability in distinguishing dangers associated with botnets. It centers basically around botnet movement in IoT networks. When consolidated, these datasets give an expansive and exhaustive reason for surveying and looking at intrusion detection techniques in Internet of Things networks.

duration	protocol_type	service	flag	wrc_bytes	dst_bytes	land	wrong_fragment	urgent	hot	...	cl_dst_svc_bin	cl_dst_sport_bin
0	0	tcp	ftp_data	SF	481	0	0	0	0	...	0	0
1	0	udp	other	SF	148	0	0	0	0	...	0	0
2	0	tcp	private	SF	0	0	0	0	0	...	0	0
3	0	tcp	http	SF	232	8153	0	0	0	...	0	0
4	0	tcp	http	SF	189	420	0	0	0	...	0	0

Fig 4 data set

id	dur	proto	service	state	spkts	dsts	dstbytes	dstbytes	rate	...	cl_dst_sport_bin	cl_dst_svc_bin	is_top_k
1	0.000011	udp	-	INT	2	0	486	0	90900.0002	...	1	2	
2	0.000008	udp	-	INT	2	0	1762	0	125000.0000	...	1	2	
3	0.000005	udp	-	INT	2	0	1988	0	200000.0059	...	1	3	
4	0.000006	udp	-	INT	2	0	900	0	180000.0008	...	1	3	
5	0.000010	udp	-	INT	2	0	2126	0	100000.0025	...	1	3	

Fig 5 data set

d) DATA PROCESSING

Data Processing

Pandas DataFrame: To work with powerful change and examination, the information is initial placed into a Pandas dataframe. This makes dealing with

the design of the dataset basic and makes an assortment of preprocessing errands simpler.

KerasDataFrame: The DataFrame might be changed into a KerasDataFrame to make it viable with Keras, considering simple reconciliation with Keras DL models.

Dropping Unwanted Columns: To streamline the dataset and improve model execution, Unwanted Columns— like IDs or pointless features— are disposed of from the DataFrame.

Visualization:

For data visualization exercises, the Seaborn and Matplotlib bundles are utilized. These libraries incorporate an assortment of visualization techniques to assist with figuring out the dissemination of information and the connections between factors, for example, disperse plots, heatmaps, and histograms.

Label Encoding:

The scikit-get familiar with library's LabelEncoder is utilized to encode categorical data. This makes it conceivable to utilize downright information in ML models by changing over clear cut names into mathematical portrayals.

The SelectPercentile strategy with Mutual Information Classification is utilized to pick features. The most valuable elements might be picked for model preparation by utilizing this measurable method, which positions features as indicated by their expectation potential about the objective variable.

e) TRAINING AND TESTING

The dataset is separated into training and testing sets to prepare and testing the ensemble-learning-based deep neural network (DNN) for assault arrangement of uneven interruption information in IoT organizations. The DNN gathering model, which utilizes numerous DNN models as base assessors, is prepared utilizing the training set. Class loads are utilized during training to address for dataset lopsided characteristics and ensure the model appropriately gains from both larger part and minority classes.

Following training, the testing set is utilized to assess the group model's precision in arranging attacks. To survey how well a model identifies attacks while decreasing false positives, execution measures including accuracy, precision, recall, F1-score, and false positive rate are registered. During testing, the ensemble DNN model's ability to oversee class imbalance in intrusion detection datasets special to IoT networks is confirmed, just like its capacity to sum up.

f) ALGORITHMS:

CNN

CNNs are deep learning architectures that examine coordinated network like information like pictures. The exploration involves CNN for IoT intrusion detection. It utilizes convolutional layers to extricate network traffic qualities and pooling layers to diminish dimensionality. Arrangement happens in completely connected layers utilizing recovered qualities. CNN's [12] ability to naturally learn progressive information portrayals makes it ideal for spotting designs in confounded network traffic, further developing IoT ID.

LSTM

Long Short-Term Memory (LSTM) recurrent neural networks (RNNs) reenact successive information with long-term conditions. The undertaking distinguishes IoT intrusions utilizing LSTM [13]. LSTM networks can dissect time-series data like network traffic better compared to feedforward neural networks since they can hold data across extended cycles. By catching transient connections in information, LSTM [13] further develops ID by detecting network action examples and abnormalities, expanding IoT network security.

CNN + LSTM

CNN+LSTM [14] is a CNN-LSTM hybrid deep learning architecture. The task involves this hybrid approach for IoT ID. CNN separates geographic data from network traffic data, while LSTM gathers transient connections. CNN+LSTM [14] further develops intrusion detection accuracy and strength by joining geological and worldly data to detect complex examples and abnormalities in network action, boosting IoT network security.

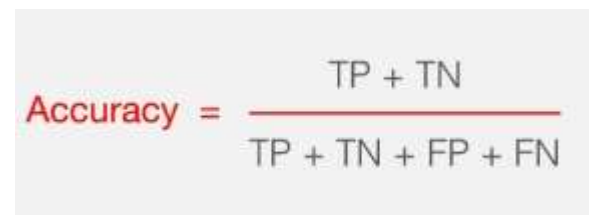
DNN

Deep Neural Networks (DNNs) incorporate various secret layers among information and result. The task involves DNN [15] as an independent IoT intrusion detection model. DNN creates progressive portrayals of approaching information to catch confounded network traffic linkages and examples. DNN[15] partitions network traffic into typical and assault classes utilizing its deep architecture, further developing IoT network security by distinguishing and moderating dangers and weaknesses.[41]

4. EXPERIMENTAL RESULTS

Accuracy: A test's accuracy is its ability to recognize debilitated from sound cases. To quantify test accuracy, figure the small part of true positive and true negative in completely broke down cases. Numerically, this is:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$



The diagram shows the accuracy formula: Accuracy = (TP + TN) / (TP + TN + FP + FN). The numerator (TP + TN) is highlighted in red, and the denominator (TP + TN + FP + FN) is in black. The entire formula is enclosed in a light gray box.

Precision: Precision estimates the level of positive cases or tests precisely sorted. Precision is determined utilizing the recipe:

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} = \frac{TP}{TP + FP}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

Recall: Machine learning recall assesses a model's ability to perceive all significant examples of a class. It shows a model's culmination in catching occasions of a class by contrasting accurately anticipated positive perceptions with complete positives.

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score: Machine learning model accuracy is estimated by F1 score. Consolidating model precision and recall scores. The accuracy

measurement estimates how frequently a model anticipated accurately all through the dataset.

$$\text{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

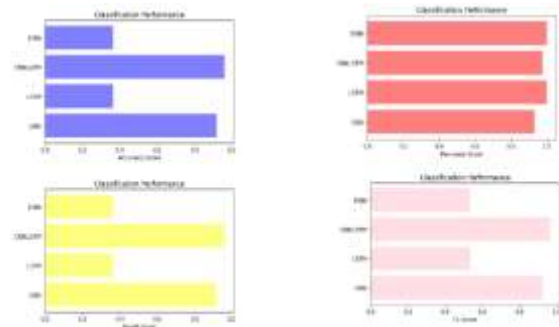


Fig 8 COMPARISON GRAPHS OF NSL KDD DATASET

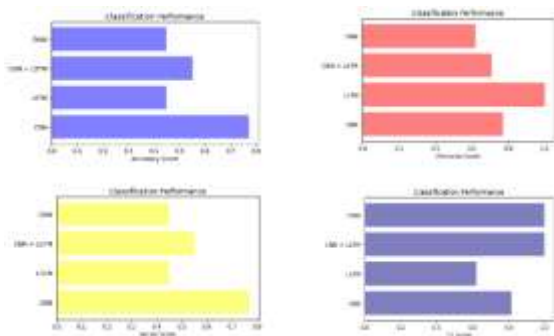


Fig 6 COMPARISON GRAPHS OF BOT-IOT DATASET

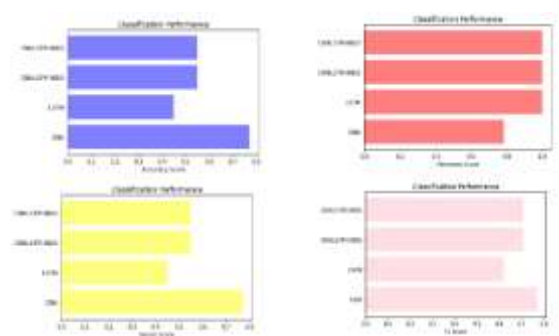


Fig 9 COMPARISON GRAPHS OF UNSW-NB15 DATASET

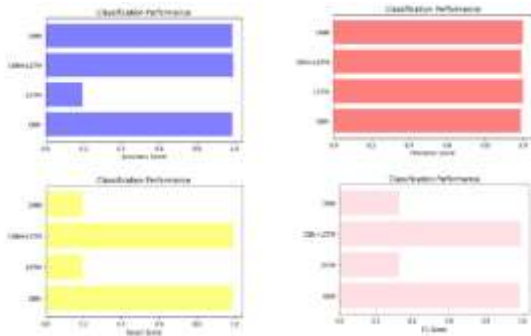


Fig 7 COMPARISON GRAPHS OF KDD-CUP DATASET

ML Model	Accuracy	f1_score	Recall	Precision
Extension CNN	0.77	0.814	0.77	0.770
LSTM	0.45	0.621	0.45	1.000
Extension CNN + LSTM	0.55	1.000	0.55	0.709
DNN	0.45	1.000	0.45	0.621

Fig 10 PERFORMANCE EVALUATION -BOT IOT DATASET

ML Model	Accuracy	f1_score	Recall	Precision
Extension CNN	0.989	0.989	0.989	0.989
LSTM	0.196	0.328	0.196	1.000
Extension CNN+LSTM	0.991	0.991	0.991	0.991
DNN	0.988	0.328	0.196	1.000

Fig 11 PERFORMANCE EVALUATION -KDD CUP DATASET

ML Model	Accuracy	f1_score	Recall	Precision
Extension CNN	0.917	0.923	0.917	0.930
LSTM	0.364	0.534	0.364	1.000
Extension CNN+LSTM	0.901	0.966	0.961	0.973
DNN	0.364	0.534	0.364	1.000

Fig 12 PERFORMANCE EVALUATION -NSL KDD DATASET

ML Model	Accuracy	f1_score	Recall	Precision
Extension CNN	0.770	0.772	0.770	0.814
LSTM	0.450	0.621	0.450	1.000
Extension CNN+LSTM	0.559	0.706	0.559	0.988
DNN	0.748	0.621	0.450	1.000

Fig 13 PERFORMANCE EVALUATION -UNSW NB15 DATASET



Fig 14 Home Page

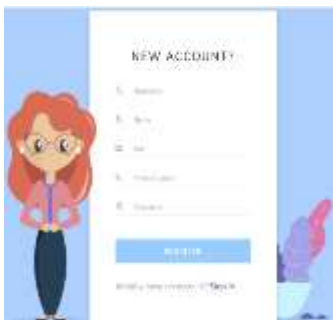


Fig 15 Sign Up



Fig 16 Sign In



Fig 17 NSL-KDD

Service

Flag

SRC Bytes

DST Bytes

Count

Fig 18 upload input data

Sensor Rate

Same SRV Rate

Diff SRV Rate

Dist Host SRV Count

Dist Host Same SRV Rate

Dist Host Diff SRV Rate

Fig 19 upload input data

Dst Host Same SRV Rate

Dst Host Diff SRV Rate

Dst Host Senior Rate

Fig 20 upload input data

Result: **There is an Attack Detected, Attack Type is DDoS!**

Fig 21 Predict result

Result: **There is an Attack Detected, Attack Type is Probe!**

Fig 22 Predict result

Result: **There is an No Attack Detected, it is Normal!**

Fig 23 Predict result



Fig 24 KDD-CUP

Protocol Type

Service

SRC Bytes

DST Bytes

Logged In

Fig 25 upload input data

Count

SRV Count

SRV Diff HOst RATE

Diff Host Count

Dst Host SRV Count

Fig 26 upload input data



Fig 27 upload input data



Fig 28 Predict result



Fig 29 Predict result

5. CONCLUSION

All in all, the bagging classifier (BC)- based deep neural network (DNN) system might tackle class unevenness in IoT intrusion detection datasets. The procedure further develops ID and arrangement across datasets by using DL and ensemble learning, as shown by promising f-score values. The utilization of DL models like CNN and hybrid CNN+LSTM further develops assault classification accuracy and strength, acquiring close to 100%

accuracy on the KDDCUP99 dataset. A Flask-based front end smoothes out testing and further develops client openness, making a functional framework interface. The review shows that the proposed procedure might tackle IoT ID issues, adding to network security.

6. FUTURE SCOPE

Assault arrangement of unequal intrusion information in IoT networks utilizing an ensemble-learning-based deep neural network incorporates a few significant features for danger recognizable proof and moderation. It utilizes feature engineering to remove geological and transient properties from network traffic information. Feature selection approaches track down helpful characteristics for fitting attack classification. Ensemble learning strategies like stowing and helping are additionally inspected to work on model execution and class lop-sidedness flexibility. The utilization of deep neural network geographies like CNNs and LSTMs grows the capacity to catch complex organization information examples and anomalies. The element scope covers all parts of making a strong ID answer for IoT networks.

REFERENCES

- [1] E. Aminanto and K. Kim, "Deep learning in intrusion detection system: An overview," in Proc. Int. Res. Conf. Eng. Technol. (IRCET), 2016, pp. 1–12.
- [2] A. Thakkar and R. Lohiya, "Attack classification using feature selection techniques: A comparative study," J. Ambient Intell. Humanized Comput., vol. 12, pp. 1249–1266, Jun. 2020.

- [3] A. Thakkar and R. Lohiya, "Role of swarm and evolutionary algorithms for intrusion detection system: A survey," *Swarm Evol. Comput.*, vol. 53, Mar. 2020, Art. no. 100631.
- [4] A. Thakkar and R. Lohiya, "Analyzing fusion of regularization techniques in the deep learning-based intrusion detection system," *Int. J. Intell.Syst.*, vol. 36, no. 12, pp. 7340–7388, 2021.
- [5] R. Lohiya and A. Thakkar, "Application domains, evaluation data sets, and research challenges of IoT: A systematic review," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 8774–8798, Jun. 2021.
- [6] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," *Future Gener. Comput. Syst.*, vol. 100, pp. 779–796, Nov. 2019.
- [7] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, MA, USA: MIT Press, 2016.
- [8] B. Dong and X. Wang, "Comparison deep learning method to traditional methods using for network intrusion detection," in *Proc. 8th IEEE Int. Conf. Commun.Softw.Netw. (ICCSN)*, 2016, pp. 581–585.
- [9] X. Liu et al., "Privacy and security issues in deep learning: A survey," *IEEE Access*, vol. 9, pp. 4566–4593, 2020.
- [10] A. Thakkar and R. Lohiya, "A review on machine learning and deep learning perspectives of IDS for IoT: Recent updates, security issues, and challenges," *Arch. Comput. Methods Eng.*, vol. 28, no. 4, pp. 3211–3243, 2021.
- [11] J. L. Leevy, T. M. Khoshgoftaar, R. A. Bauder, and N. Seliya, "A survey on addressing high-class imbalance in big data," *J. Big Data*, vol. 5, no. 1, pp. 1–30, 2018.
- [12] A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary," *J. Artif. Intell. Res.*, vol. 61, pp. 863–905, Jan. 2018.
- [13] D. Elreedy and A. F. Atiya, "A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance," *Inf. Sci.*, vol. 505, pp. 32–64, Dec. 2019.
- [14] S. Bagui and K. Li, "Resampling imbalanced data for network intrusion detection datasets," *J. Big Data*, vol. 8, no. 1, pp. 1–41, 2021.
- [15] B. Mirza, Z. Lin, and K.-A. Toh, "Weighted online sequential extreme learning machine for class imbalance learning," *Neural Process.Lett.*, vol. 38, no. 3, pp. 465–486, 2013.
- [16] P. Bedi, N. Gupta, and V. Jindal, "Siam-IDS: Handling class imbalance problem in intrusion detection systems using siamese neural network," *ProcediaComput. Sci.*, vol. 171, pp. 780–789, Jun. 2020.
- [17] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches," *IEEE Trans. Syst., Man, Cybern.C,Appl.Rev.*, vol. 42, no. 4, pp. 463–484, Jul. 2012.

- [18] E. Hodo, X. Bellekens, A. Hamilton, C. Tachtatzis, and R. Atkinson, “Shallow and deep networks intrusion detection system: A taxonomy and survey,” 2017, arXiv:1701.02145.
- [19] C. Sun, K. Lv, C. Hu, and H. Xie, “A double-layer detection and classification approach for network attacks,” in Proc. 27th Int. Conf. Comput. Commun. Netw. (ICCCN), 2018, pp. 1–8.
- [20] Y. Yuan, L. Huo, and D. Hogrefe, “Two layers multi-class detection method for network intrusion detection system,” in Proc. IEEE Symp. Comput. Commun. (ISCC), 2017, pp. 767–772.
- [21] R. Abdulhammed, M. Faezipour, A. Abuzneid, and A. AbuMallouh, “Deep and machine learning approaches for anomaly-based intrusion detection of imbalanced network traffic,” IEEE Sens. Lett., vol. 3, no. 1, pp. 1–4, Jan. 2019.
- [22] J. Jiang, Q. Wang, Z. Shi, B. Lv, and B. Qi, “RST-RF: A hybrid model based on rough set theory and random forest for network intrusion detection,” in Proc. 2nd Int. Conf. Cryptogr. Security Privacy, 2018, pp. 77–81.
- [23] L. Breiman, “Bagging predictors,” Mach. Learn., vol. 24, no. 2, pp. 123–140, 1996.
- [24] S. Hido, H. Kashima, and Y. Takahashi, “Roughly balanced bagging for imbalanced data,” Stat. Anal. Data Min. ASA Data Sci. J., vol. 2, nos. 5–6, pp. 412–426, 2009.
- [25] A. Kadiyala and A. Kumar, “Applications of python to evaluate the performance of bagging methods,” Environ. Progr. Sustain. Energy, vol. 37, no. 5, pp. 1555–1559, 2018.
- [26] B. Ghogh and M. Crowley, “The theory behind overfitting, cross validation, regularization, bagging, and boosting: Tutorial,” 2019, arXiv:1905.12787.
- [27] R. Lohiya and A. Thakkar, “Intrusion detection using deep neural network with antirectifier layer,” in Applied Soft Computing and Communication Networks. Singapore: Springer, 2021, pp. 89–105.
- [28] A. Thakkar and R. Lohiya, “A review of the advancement in intrusion detection datasets,” Procedia Comput. Sci., vol. 167, pp. 636–645, Apr. 2020.
- [29] M. Mulyanto, M. Faisal, S. W. Prakosa, and J.-S. Leu, “Effectiveness of focal loss for minority classification in network intrusion detection systems,” Symmetry, vol. 13, no. 1, p. 4, 2021.
- [30] G. Kyriakides and K. G. Margaritis, Hands-On Ensemble Learning with Python: Build Highly Optimized Ensemble Machine Learning Models Using Scikit-Learn and Keras. Birmingham, U.K.: Packt Publ. Ltd., 2019.
- [31] A. Thakkar and R. Lohiya, “A survey on intrusion detection system: Feature selection, model, performance measures, application perspective, challenges, and future research directions,” Artif. Intell. Rev., vol. 55, no. 1, pp. 453–563, 2022.
- [32] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic minority over-sampling technique,” J. Artif. Intell. Res., vol. 16, pp. 321–357, Jun. 2002.
- Dataset: links
- NSL - KDD :

<https://www.kaggle.com/datasets/kaggleprollc/nsl-kdd99-dataset>

KDD-CUP :

<https://www.kaggle.com/datasets/galaxyh/kdd-cup-1999-data>

UNSW-15-NB:

<https://www.kaggle.com/datasets/sweety18/unsw-nb15-training>

Bot-IoT

[:https://www.kaggle.com/datasets/vigneshvenkateswaran/bot-iot-5-data](https://www.kaggle.com/datasets/vigneshvenkateswaran/bot-iot-5-data)

[33] G.Viswanath, “Hybrid encryption framework for securing big data storage in multi-cloud environment”, Evolutionary intelligence, vol.14, 2021, pp.691-698.

[34] Viswanath Gudditi, “Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage”, Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol.12, 2021, pp.545-552.

[35] Viswanath Gudditi, “A Smart Recommendation System for Medicine using Intelligent NLP Techniques”, 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), 2022, pp.1081-1084.

[36] G.Viswanath, “Enhancing power unbiased cooperative media access control protocol in manets”, International Journal of Engineering Inventions, 2014, vol.4, pp.8-12.

[37] Viswanath G, “A Hybrid Particle Swarm Optimization and C4.5 for Network Intrusion Detection and Prevention System”, 2024, International Journal of Computing, DOI:

<https://doi.org/10.47839/ijc.23.1.3442>, vol.23, 2024, pp.109-115.

[38] G.Viswanath, “A Real Time online Food Ordering application based DJANGO Restfull Framework”, Juni Khyat, vol.13, 2023, pp.154-162.

[39] Gudditi Viswanath, “Distributed Utility-Based Energy Efficient Cooperative Medium Access Control in MANETS”, 2014, International Journal of Engineering Inventions, vol.4, pp.08-12.

[40] G.Viswanath,“ A Real-Time Video Based Vehicle Classification, Detection And Counting System”, 2023, Industrial Engineering Journal, vol.52, pp.474-480.

[41] G.Viswanath, “A Real- Time Case Scenario Based On Url Phishing Detection Through Login Urls ”, 2023, Material Science Technology, vol.22, pp.103-108.

[42] Manmohan Singh,Susheel Kumar Tiwari, G. Swapna, Kirti Verma, Vikas Prasad, Vinod Patidar, Dharmendra Sharma and Hemant Mewada, “A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification” published in Journal of Computer Science, Available at: <https://pdfs.semanticscholar.org/69ac/f07f2e756b79181e4f1e75f9e0f275a56b8e.pdf>