A PROJECT REPORT

ON

"NETWORK INTRUSION DETECTION SYSTEM BASED ON GAN"

REPORT SUBMITTED TOWARDS PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE OF

BACHELOR OF TECHNOLOGY IN ELECTRONICS AND TELECOMMUNICATIONS ENGINEERING WITH SPECIALIZATION
IN
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted By

Students Name	PRN
DHRUV DOLAS	**20070123015**
VARUN SINGH	**20070123050**
VAMSI KRISHNA. S	**20070123059**
AALIYA SHAIKH	**20070123084**

UNDER THE GUIDANCE OF

Dr. Sashikala Mishra

Professor

॥वसुधैव कुटुम्बकम्॥

SYMBIOSIS INSTITUTE OF TECHNOLOGY

SYMBIOSIS INTERNATIONAL (DEEMED UNIVERSITY)

Pune - 412115

2020-2024

TITLE OF THE PROJECT

REPORT SUBMITTED TOWARDS PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE OF

BACHELOR OF TECHNOLOGY IN ELECTRONICS AND TELECOMMUNICATIONS ENGINEERING WITH SPECIALIZATION
IN
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

By

DHRUV DOLAS	**20070123015**
VARUN SINGH	**20070123050**
VAMSI KRISHNA. S	**20070123059**
AALIYA SHAIKH	**20070123084**

Under the guidance of

SASHIKALA MISHRA
Professor

CERTIFICATE

This is to certify that the Project work entitled “**Networks Intrusion Intrusion Detection Systems Based on GAN**” is carried out by Dhruv Dolas**,** Varun Singh, Vamsi Krishna S a Aaliya Shaikh in partial fulfillment for the award of the degree of **Bachelor of Technology in Electronics and Telecommunications** with Specialization in (AIML) at Symbiosis Institute of Technology Pune, Symbiosis International (Deemed University) Pune, India during the academic year 2023-2024.

-------------	---------	---------
Dr. Deepali Vora	Dr. Sashikala Mishra	Director

SYMBIOSIS INSTITUTE OF TECHNOLOGY
SYMBIOSIS INTERNATIONAL (DEEMED UNIVERSITY)
Pune - 412115
2020-2024

DECLARATION

We declare that the project titled "**Network Intrusion Detection**" has been submitted to Symbiosis Institute of Technology Pune, Constituent of Symbiosis International (Deemed University) Pune for the award of the degree of Bachelor of Technology with Specialization in AIML is a result of original project work has been carried out in this thesis. It is further declared that the project report or any part has not been previously submitted to any University or Institute for the award of a degree.

Name of Student(s) : Dhruv Dolas, Varun Singh, Vamsi Krishna, Aaliya Shaikh

PRN : 20070123015,20070123054,20070123059,20070123084

Degree : Bachelor of Technology in Electronics and Telecommunications

Specialization : Artificial Intelligence and Machine Learning

Title of the project : Network Intrusion Detection System Based on GAN

(Name and Signature)

Date:

ACKNOWLEDGEMENT

We want to express my sincere gratitude to everyone who supported us throughout this project. First and foremost, we would like express one's thanks to our project guide, **Dr. Sashikala Mishra**, for her valuable guidance, encouragement, and feedback. She has been a constant source of inspiration and motivation for me.

We would also like to express gratitude to head of the department, **Dr. Deepali R. Vora**, for providing us with the necessary facilities and resources for conducting this project. We are grateful to her for her constant support and advice.

We want to acknowledge the contribution of our project team members, **Dhruv Dolas, Varun Singh, Vamsi Krishna.S, Aaliya Shaikh**, who have worked hard and cooperated with in every project stage. They have been accommodating and supportive throughout this journey.

We would also like to express one's thanks to our family and friends for their support, care, and love. They have always been there for me in times of need and stress. They have encouraged me to pursue my passion and achieve my goals.

Lastly, I thank **Symbiosis Institute of Technology Pune** for allowing me to work on this project and enhance my skills and knowledge. We are proud to be a part of this prestigious institution.

ABSTRACT

In today’s globally connected digital environment, network security stands as an ever-pressing concern, given the constant evolution of cyber threats. In response to this dynamic landscape, the integration of generative adversarial networks (GANs) into network intrusion detection systems (NIDS) has emerged as a groundbreaking strategy. Traditional intrusion detection techniques, which primarily rely on predetermined rules and signatures, often struggle to identify innovative and complex threats that deviate from established patterns. These conventional systems are inherently limited by their rigidity, as they are less adaptable to emerging threats. GAN-based intrusion detection systems (GAN-IDS), on the other hand, present a promising approach to address these limitations. They operate by learning from historical data and continuously adapting to new and evolving threats, allowing organizations to bolster their network security in a proactive and dynamic manner. This study aims to explore the impact of GAN-IDS on key performance metrics such as detection precision, false-positive rates, and its ability to recognize zero-day assaults, shedding light on the potential benefits of this innovative approach in the realm of network security. The incorporation of GANs into NIDS has the potential to significantly enhance detection precision by enabling systems to discern even subtle deviations from normal network behavior. This increased precision can lead to more accurate threat identification, reducing the risk of overlooking malicious activities. Moreover, GAN-IDS systems are designed to adapt and evolve alongside the ever-changing threat landscape, ultimately lowering false-positive rates. By minimizing the occurrence of false alarms, organizations can allocate their resources more efficiently, focusing their attention on genuine threats rather than spending time investigating benign anomalies. Furthermore, the ability of GAN-IDS to recognize zero-day assaults, which are previously unseen and unclassified threats, can be a game- changer in network security. Traditional NIDS often struggle with zero-day attacks, but GAN-IDS, through their ability to learn from historical data and adapt to new patterns, offer a robust defense against these novel and potentially devastating threats.

TABLE OF CONTENTS

Certificate

Declaration

Acknowledgment i

Abstract ii

Table of Contents iii

List of Tables v

List of Figures vi

List of Abbreviations ix

CHAPTER 1: INTRODUCTION 8

1.1 Introduction 8

1.2 Problem statement 9

1.3 Scope of research 9

1.4 Research hypothesis 9

CHAPTER 2: LITERATURE REVIEW 10

2.1 Background 10

2.2 Summary of literature review and research gap 11

CHAPTER 3: SOFTWARE REQUIREMENTS SPECIFICATION 23

CHAPTER 4: METHODOLOGY 25

CHAPTER 5: IMPLEMENTATION 38

CHAPTER 6: RESULTS 43

45 **CHAPTER 7: CONCLUSION** 45

REFERENCES

APPENDICES

Plagiarism Report

LIST OF TABLES

	Table name	Page No
1	Essential Software Components Table	25
2	Comparison of Key Characteristics Across Network Intrusion Detection Datasets.	31
3	Performance of the model on different sets of data	44

LIST OF FIGURES

	Figure name	Page No
1	GAN based Architecture	29
2	Sequential flow of diagram of the process	34
3	Denial of Service Loss and Accuracy	35
4	Results for R2L Loss and Accuracy	35
5	Probe Loss and Accuracy	36
6	Results for U2R Loss and Accuracy	36
7	Processing Diagram for NSS KDD Dataset	42

LIST OF ABBREVIATIONS

Abbreviated Word	**Expansion**
GAN	Generative Adversarial Network
NID	Network Intrusion Detection
IDS	Intrusion Detection System
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
DNN	Deep Neural Network
ML	Machine Learning
DL	Deep Learning
DoS	Denial of Service
DDoS	Distributed Denial of Service
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
IP	Internet Protocol
API	Application Programming Interface

INTRODUCTION

1.1 Introduction

In today's intertwined and vastly proliferating digital landscape, network security has proven to be of paramount importance. The relentless and ever innovating threats that await actually demand continuous advancements in the field of intrusion detection. Regular security solutions, which are predominantly dependent on predefined rules and signatures are increasingly challenged by the adaptability of the modern attacking strategies [1]. To address these formidable challenges and ensure that there is a certain degree of resilience of network defenses, our project introduces a comprehensive and innovative approach which includes the fair integration of GANs into the detection systems of network intrusions[3].

The generative adversarial networks which are known for their capability to capture intricate data ditributions and generate realistic data models happen to certainly hold promising potentiality to bring forth a revolution in the accuracy and adaptibility of intrusion detection techniques[4]. The ultimate and the core objective of this study is to delve deeply into the GAN IDS model, thereby offering a comprehensive understanding of its implications on essential metrics of intrusion detection performance. By the power of machine learning and its subset deep learning, GAN based Intrusion Detection Systems aspire to meticulously analyze intricate the network infrastucture patterns, dstrike a distinction between legitimate and malicious activities, and even unveil the presence of a zero day attacks threats that were previously oblivious to the cybersecurity teams[5].

The vitally important need for the state-of-the-art technologies such as the GAN based IDS has never been more pronounced, as current threat landscape evolves at an unprecedented and unparalleled pace [6]. This particular study happens to contribute vastly to the ongoing cybersecurity discourse by offering empirical insights into the effectiveness of GAN IDS [6]. It seeks to address critical concerns such as reducing false positive rates, enhancing detection precision and fortifying the system's resistance to an array of threats that happen to await [7]. Furthermore, our study aims to provide deeper appreciation of the transformative role that is played by the GAN IDS in emboldening and bolstering the network security within an era defined by escalating and ever evolving cyber threats[8].

On subject to vast investigation, not only the potential advantages but rather also the challenges associated with this particular approach, our research endeavors to offer a comprehensive guidance to entities looking forth to enhance their network security and the resilience posture[9]. As the digital landscape continues to advance, the integration of the GANs into the NIDS portrays a significant stride toward proactive and adaptable cybersecurity solutions, ultimately empowering institutions to stay one step ahead of the dynamic threat landscape[10].

1.2 Problem Statement

The primary challenge addressed in this research revolves around the need for more sophisticated and robust methods in network intrusion detection systems (NIDS). Traditional approaches might struggle to keep pace with evolving and sophisticated cyber threats, necessitating innovative techniques for effective intrusion detection and response.

1.3 Scope of Research

The research aims to explore the utilization of Generative Adversarial Networks (GANs) in network intrusion detection. This includes the application of GANs to generate synthetic data closely resembling real network traffic, thereby facilitating the development of more efficient intrusion detection models.

1.4 Research Hypothesis

The hypothesis underlying this research is that the implementation of GANs in network intrusion detection systems will enhance the accuracy and robustness of identifying and mitigating security breaches in network traffic.

1.5 Objectives

The key objectives of this research are as follows:

Investigate the feasibility of GANs in generating synthetic data representative of network traffic patterns.

Develop and evaluate Network Intrusion Detection Systems (NIDS) based on GAN-generated synthetic data.

Assess the effectiveness of GAN-based NIDS in identifying and responding to various intrusion attempt

LITERATURE SURVEY

2.1 Background Study

The research landscape surrounding network intrusion detection systems (NIDS) is vast, encompassing a variety of techniques and strategies aimed at strengthening cyber defense against a variety of evolving threats New to improve the performance, accuracy and changes have increased -Various experts and research groups exploring approaches have contributed to this work Report NIDS faces challenges in effectively dealing with known and unknown threats, prompting exploration of sophisticated techniques such known as deep learning and machine model to improve input detection capabilities

Several studies such as Mukherjee et al [1], Raghunath et al [2], Jiang Ke et al [3] have delved into intrusion detection by various techniques and have highlighted the importance of machine learning and data mining techniques use emphasis in order to detect threats and anomalies in network traffic. Notably, the convergence of deep hierarchical networks and hybrid sampling by Jiang Ke et al., and the promising performance of the resulting classifier demonstrate the potential of deep learning to provide intrusion detection capabilities has been great.

Furthermore, Muhammad Ashfaq Khan's research [4] successfully addressed the issues of gradient problem and computational complexity in Network Intrusion Detection Systems (NIDS), and developed a model capable of analyzing spatial-temporal reliability for intrusion detection Farrukh Aslam Khan et al, autoencoder with dual. Using complex neural networks based on comprehensive deep learning models, we obtained impressive accuracy demonstrating the power of sophisticated neural networks in NIDS.

Moreover, Adele Binbusayis et al [6] introduced an unsupervised learning method for intrusion detection, and emphasized the importance of improving the relationship between classification feature representation and learning, thus moving accuracy effectiveness Thasin et al [7] tested different automation algorithms, random in detecting malicious packets -Determining the effectiveness of the forest process, set up for further research on deep learning algorithms

Other developments include Leslie et al.'s model of adversarial attack [9], Salem et al.'s analysis of GANs for anomaly detection [10], and Park et al.'s AI-based NIDS that handles unweighted data want to address challenges and Chhetri et al. based model and the integration of Yin-based GAN in botnet detection exemplifies the types and applications of GANs in different security domains

However, the research project is not without its challenges. Studies on the susceptibility of GAN models to black-box attacks by Zhao et al [14] and synthetic flow-oriented traffic generators by Ring et al [13] highlight the need for further research and emphasizing the innovative solutions to such weaknesses

Accumulated research indicates a gradual shift towards the use of advanced machine learning, deep learning, and GAN-based techniques to enhance intrusion recognition systems All these studies show interest a there is a growing number of new strategies aimed at addressing dynamic and multifaceted cybersecurity threats and highlights the growing importance in this area The study suggests that future developments and promising developments in monitoring network security.

2.2 Summary of literature review and research gap

As larceny of information can result in significant loss, protecting computer and network information for organizations and individuals has become a crucial undertaking. To avoid this, intrusion detection systems are utilized. As a way to improve IDS's functionality, various machines learning strategies are created. The primary goal is to address the issue of intrusion's adaptability. The suggested IDS includes the ability to identify the most common attacks as well as unidentified attacks. The suggested IDS as per [8] includes three key fundamental modus operandis: Decision Maker, Clustering Manager and Update Manager. In particular the NSL-KDD data set has been subjected to prediction followed by forecasting how well has the suggested IDS would literally function. Both the use of supervised and unsupervised methods are accompanied in order to have an optimized and better efficacy.

Mukherjee et al [1] have successfully performed their research and vouched for the fact that Intrusion Detection happens to be a feasible and most adaptable form of security to implement in the cyber world. Since the basis is on host audit trail and network, the study of traffic patterns and the analysis becomes pretty simpler as opposed to using other methods. They have performed accurate research with novel strategies and have used benchmark mechanisms in order to successfully implement intrusion detection systems in the networks.

Raghunath et al [2] have researched and proposed strategies to implement network intrusion detection in a sophisticated manner. They have used data mining suite techniques to automatically detect threats that await the networks of computers. They have implemented an unsupervised detection strategy in order seek a better safety score. Their emperical results certainly show a fair amount of resilience to the threats. This particular research has laid a foundation for future researchers to develop optimized techniques in order to have the networks safeguarded from all sorts of dangers.

Jiang K et al[3] have proposed a multitude of strategies which certainly involve incorporation of a vivid range of measures in order to mitigate the risks and threats that await the networks. Primarily a particular method that is based on the amalgamation of two key strategies which are the deep hierarchical networks and the hybrid sampling have been employed. They have trained their model by establishing a fair balance amongst their dataset. This was essential done in order to decrement the training time significantly. They have also enmeshed a different ploy to counter the multidimensional cyber threats that are apt for the deep hierarchical networks. The generated model happened to be in a state where the it could extract features by taking the utmost advantage that pertains to machine learning subfield of deep learning. This classifier which has been created has shown results that were absolutely promising and it has outhustle the conventional ML algorithms such as the random forest , AlexNet and the LeNet.

Muhammad Ashfaq Khan in his research [4] has executed a NIDS that was based on HCRNN. In their study they have attempted and were successful in countering the trouble of vanishing gradient to large extent. The issue of exploding gradient was also addressed. The created NID model was competent enough to analyse the robust features pertaining to the robust features of both the spatial and temporal dependencies. The ultimate agenda was to have the best of both the worlds that is to club both the anomaly based and the signature-based methods to decrement the computational complexity and spike the accuracy and DR for the purpose of detecting the intrusion.

From their study Farrukh Aslam Khan et al [5] have conjectured and later developed a model which is based on the dual stage DL model which is based again on a deep and profoundly stacked neural network based on autoencoder. The dual stage process fundamentally encompasses the a double hidden layer comprising of a classifying feature based on softmax function. The training is done with partial supervision. Each layer that is hidden is trained and exposed to a large set of large data and its features which do not possess any labels. They have also made use of the datasets which are publicly available and are considered yardsticks in the domain. The ultimate model which was developed for the sake of network intrusion detection has exhibited and astounding accuracy of 99.996 percent for the KDD99 Dataset. On the other hand, it has developed an accuracy of 89.134 percent for the UNSW- NB15 dataset. They purport to work forth and develop a multi task adaptable algorithm in the future to optimize the model for detecting the network intrusions.

Adel Binbusayyis et al [6] have developed a model that implements a fairly sagacious method that involves an approach based on unsupervised learning for intrusion detection of networks. The key features of their work certainly happen to be the establishment of a pathway between the feature representation and the learning of the classifier which happens to be a feature that is bereft and devoid in the conventional intrusion detection systems. The second salient feature that was designed was the ability of the model in order to seek forth and learn from the data of the network traffic. Henceforth resulting in heightened magnitude of the accuracy as opposed to the pre existing conventional intrusion detection system models.

Thaseen et al [7] in their research have extensively undertaken an exercise to execute the testing of a multitude of automations algorithms such as the random forest, naive bayes, svm including knn. By subjecting all these models to rigorous scrutiny followed by testing the model on popular datasets they have found that the random forest algorithm was the most efficient algorithm in detecting the packets which are malicious. Ultimately the have also discerned the multi-faceted attributes of the data packets. A wise distinction between a normal packet , an encrypted one and an encrypted malicious one was was stricken by the model and they have vouched for the fact that this was vastly due to the implementation of random forests in comparison to the other algorithms.

Subsequently Deep learning algorithms had been moulded too for detection of perpetration and the recognition of packets which is encompassed by the future scope.

In their work Leslie et al[9] have created a model for an adversarial attack method to generate an adversary which in contrast to a black box intrusion identification model. The synthesized feature points by them were not just confined to attack the black box but have also been employed to generate taxonomical samples that pertain to the unlabeled ones. For deploying this, they have used both the GAN and the conventional Intrusion Detection model in attempt to restrict the operation within the bounds and fetters of the feature space. By this means they were in a position where they could prognosticate the adversarial attack based out of the selection of features of the GAN model. An exercise was undertaken to acutely increase the precision of the model by reshaping the malicious traffic using wireless communications on the grounds of the computed values.

Salem et al [10] have worked vastly on the grounds of host-based intrusion detections alongside generation of certain anomalies using GANs. In their paper, they state about their foremost trail to incorporate Cyclical GAN to synthesize some sort of anomalous data from the pre existing regular data. The adfa-ld dataset's few images were fed into the cycle GAN. The fabricated aberrations have been subsequently amalgamated with the actual dataset and were subjected to a particular MLP. The results broadly exhibited something promising and those were astoundingly showing the increased percentage of unseen anomaly detection from roughly 17 percent to a vast 80 percent. All in all this form of approach had outperformed the pre existing SMOTE, highlighting the salience of GANs.

Idris et al [9] in their paper have introduced "EdgeIDS," an Intrusion Detection System (IDS) for IoT devices using the Skip-GAN anomaly generative adversarial network. 15 "EdgeIDS" outperforms current methods, as found quantitatively. Their experimental results highlight the effectiveness in detecting traffic and indirect attacks in IoT environments. Because of the limited capacity of most IoT devices, "EdgeIDS" focuses on analyzing inbound network data in real time. In addition, "DL-NIDS", another network IDS, increases security by deployment on fog nodes, especially for devices with higher processing power, which is best supported by small hardware acceleration.

As an illustration for NID, Chhetri et al. [12] have presented a model based on conditional GANs to satisfy essential security requirements by closely examining the intrinsic and complex interactions among the cyber as well as physical realms within the CPS. On integrating the GANs with a Botnet detection model, Yin et al. constructed a framework that improves the efficacy of spotting dangerous attacks while maintaining the fundamental qualities of the actual detection model intact. A particular GAN-based model contributed by Seo et al. was created to lower the false positive rate that pertains to the networks of the vehicle, enhancing the safety of driver.

In their research Ring et al [13] they have conjectured and put forth three synthetic flow oriented network traffic generators using the time scale rule of updation. The first approach included the IP addresses and ports as continuous packets in a particular interval of 0 and 1. The second approach incorporated a plan where binary attributes from categorical data and the numerical data were created on Bwgangp. The third approach was designed in such a way that the model is able to learn meaningfully based on continuous categorical representation of attributes such as that of the internet protocol addresses etc.

In their research study Zhao et [14] al have put forth a GAN model for the attack which is against the black box intrusion. The final outcome and the experimental results pertaining to their experimentation have ensured the function of the traffic instances. The synthesized adversarial samples have successfully escaped the detection protocol of the Intrusion Detection. This model has outhustled the pre-existing algorithms in most of the ways.

In their work Zixu et al [15] have presented an unsupervised model for anomaly detection approach using the tool of Generative Adversarial Networks for networks that have been distributed for Internet of things. The have subjected their model to the yardstick dataset which is the dataset of University of New South Wales- Iot, thereby having attained an optimized performance of the model with enhanced resilience to the threats that are around.

Shuokang Huang et al have [16] introduced a detection framework for the intrusion system in order to tackle the challenge of imbalancement of class in the data that would hinder the performance of many intrusion identification systems. Imbalanced GAN (IGAN) model was used and compared against other conventional methods. IGAN produced a 84\% accuracy as compared to w/o IGAN (79\%).

Eunbi Seo et al.[17] Developed an Intrusion Detection System for In Vehicle Network using GAN in for reducing risks of the vehicle from all the ECUs and subsequently relevant provided in the vehicle from the threats that await. Due to limited targets since only some threats are reflected when constructing a system, a GIDS (Generative Adversarial Nets based Intrusion Detection System) was used to solve these problems. Each of the four attacks (DOS,FUZZY<RPM,GEAR) was have been detected with an average accuracy of about 98 percent.

Mohiuddin et al. [18] brought in the method of Random Forest (RF) for automated detection of intrusions to demonstrate the superiority it is to its predecessor which happens to be the Support Vector Machine (SVM) which has been vividly used in IDS. The detection rate can reach 65\% when the false positive is 1%.

Saleh et al[19]. has developed a detection system for intrusion which is named to be Least Square Support Vector Machine based IDS (LSSVM-IDS) constructed from the features that have been selected on the basis of algorithm using KDD Cup 99, NSL-KDD and Kyoto 2006+ dataset. It produced an accuracy for the classes Normal, Denial of Service, the Probe and R2L with 89.31%, 99.27%, 84.16% and 48.13%, respectively.

Tang et al [20] state that deep learning plays an essential role in SDN flow-based anomaly detection. The recent acknowledgment and adoption of SDN as a viable network management solution emphasize its relevance with regards to deep neural networks detecting anomalies within reasonable time limits. Some deep learning approaches to improve network security with SDNs as a component of the future Internet architecture of tomorrows using only some of these features extracted from the NSL-KDD Dataset set.

Zhu et al [21] tackle the problem of low detection accuracy of classic PSRNIDSs and propose a new way to improve their efficiency. This approach relies on the random forest algorithm to enhance power system network intrusion detection. The approach involves several key steps: building a random forest decision tree by selecting a power system network intrusion sub-sample, optimizing the random forest model using an edge function, assessing the vector accuracy through the minimum state vector of the power system network, estimating the measurement residual for attack detection in power system networks, and It is an efficient and complete method for detecting intrusions of a power system network.

Dlamini & Fahim's study [22] concentrates on improving anomaly detection through better representation of minor classes to lead to enhance classifier performance. To produce false minority samples, they utilize domain-conditional generative adversarial networks under DGM (unbalanced data set). Using a combination of GANs, KL-divergence based learning helps in adequately characterizing complex minorities class distribution revealing its importance.

Pallavi et al. [23] Pallavi and some other researchers did a study to see if using convolutional neural networks (CNNs) could help catch hackers trying to sneak into interactive systems better than other machine learning methods. They tried out this special CNN they designed on some fake hacking attempts with 10 people using Jupyter Notebook. They ran the numbers to figure out how accurate their CNN was at catching the fake hacks compared to a regular artificial neural network (ANN). In the discussion part, they talked about how they analyzed the predictions from the ANN to see how well it could pick out hacking tries in the fake student interactive systems they set up. The ANN was 93% accurate according to their stats. The CNN did a little worse at 91% accuracy and still, the results were pretty reliable overall, with a 95% confidence interval. So in the end the CNN performed better than the ANN at predicting hacks in the mock student interactive systems. This suggests CNNs might be the way to go for improving hack detection in real interactive systems. The study made a good case for CNNs doing this job better than other machine learning approaches.

A study conducted by Hinton and Zemel [24] on auto-encoders and MDL principles has considerable practical relevance for enhancing the information processing method and training of NIDS. The emphasis is on the application of autoencoders, which are used for efficient data description through decreasing both the value and size of a single code vector as well as a reconstruction error at once. Lastly, it focuses on how auto-encoders can provide simple but very powerful IDSs capable of detecting complicated data patterns existing in complex networks.

Samrin and Vasumathi [25] "Deep-diving into the significant arena of intrusion detection in computer systems and networks, Samrin and Vasumathi [8] bring to fore a key aspect crucial to cybersecurity. The dynamic nature of new attack forms often pose considerable challenges for traditional defenses. They emphasize on Intrusion Detection Systems (IDS) broken down primarily as Network-based IDS and Host IDS (HIDS). These setups play pivotal roles in shielding network operations from harm's way correlated with destructive activities. One persistent problem articulated by them is the tedious workload associated with manual sifting through volumes of network data, which ultimately leads to resource exhaustion both monetarily & timewise. Their work proposes an innovative solution: leverage data mining techniques that can distill valuable patterns within massive datasets relating particularly around system intrusions.

Kim and Pak realm ebbing towards early recognition methods vis-à-vis potential threats triggered within secured networks [26]. Unlike prior means dependent wholly upon session-heavy information retrieval; they underscore more real-time tangible packet-data observance focused at detecting possibly damaging traffic journey 15 flowing invisibly across supposedly secure realms subjectively flagged green otherwise.{New paragraph}In order bypass false-positive flags so recognized it suggests applying Generative Adversarial Networks(GAN)s designed consciously aimed align improved intrusive classification foresight by learning disregard unimportant packets irrelevantly presenting themselves amidst sea inter-network transactions vitally occurring unnoticed underneath casual viewers line sight besides incorporating LSTM-DNN precisely drawing best possible versions accuracy plus speed thereby outlasting existing models paths tread earlier identified experiments prove their claimed strength thus managing expose mischievous actors lurking unperturbed behind firewalls causing silent chaos."

In relation to anomaly detection in CPS systems, Li et al.’s work [27] is concerned with exploiting GANs. The study notes that GAN based anomaly detection can be a successful way of detecting unusual behaviors in huge networked systems like by using LSTM-RNN architectural techniques. The study shows that GAN-AD is efficient in detecting SWAT system through the process detection using CISA model. Additionally, the paper proves that GAN could be applied for the modeling purposes in predicting the future multivariate time series distribution. In general, evaluating all the different research contributions emphasizes more on using modern machine learning algorithms and AI based techniques for designing effective and strong NIDS. The overall essence of these studies is that one needs to use hybrids learning methods, machine learning algorithms, generative models or advanced data analytics techniques in order to adequately handle the emerging issues of network security as well as detecting anomalies in complex technological environments. Thus, innovative and adaptive IDS should be applied for protecting modern critical network infrastructures.

Liu and Lang [28] perform an extensive review of IDS based on machine and deep learning addressing networks in contemporary society that require strong cybersecurity mechanisms. The survey provides a structured vocabulary with relation to IDS, grouping and concising the related knowledge objects by data items. This outlines the impact of machine-learning algorithms in improving detection efficiency, reducing false alarms, and detecting unidentified attacks. The authors outline a strong basis for cyber security researchers through their discussion on deep learning performance as well as multitasking capabilities Survey provides comprehensive description of core terminologies, taxonomy, methodology, metric and benchmark dataset with practical considerations using ML/DL for tackling relevant security problems in IDS. It also provides an overview of the current challenges and future perspectives as well as a critical review of the most notable studies conducted recently.

As demonstrated by Imamverdiyev and coauthors [29] and presented by Abdullayeva, there is a thorough research that examines how the deep learning methods can be applied in order to detect the Denial of Service attack with Gaussian-Bernoulli RBM model The proposed deep RBM model attains greater accuracy for detecting DoS attacks by adding seven extra layers between the visible and hidden layers. Using the perpetual, data amicable form of RBM, whereby the visible layer bears a pattern of a 16 Gaussian, the authors perform a contrastive assessment with other deep learning schemes, such as Bernoulli-Bernoulli RBM and deep belief networks. The assessment of this NSL-KDD dataset shows that the proposed multilayer deep Gaussian-Bernoulli type RBM performs best and can hence be used as a reliable detector of DoS attacks (Imamverdiyev et al., 2018).

Hashem Haghighat et al [30] and Jun Li presented a new deep leaning architecture based on voting called VNN for intrusion detection that strives to solve the never-ending problem of successful attacks in computerized networks. Using different deep leaning architecture, the VNN framework aggregates various number of models increase detection accuracy and resilient against complexity security attacks like DDoS, honeypot and phishing. Using experiments with datasets that are generally accepted, such as the KDDCUP'99 and CTU-13, the study shows the ability of the voting algorithm dramatically reduces a false alarm rate up to 75% against the initial deep learning models consisting of for instance DNN, This study is an important step towards enhancing the performance of intrusion detection and it assists security experts in countering sophisticated cyber-attacks.

In 2023, Lee et al. [31] introduce a new AI-based technique for detecting cyber threats, which includes the application of artificial neural nets in an AI-SIEM scheme. They use event profiling as a pre-processing technique combined with different neural network model such as FCNN, CNN and LSTM to discriminate between true positives and false positives. They thoroughly experiment with benchmark and actual data sets using their proposed model, showing how it outshines many existing machine learning approaches. It can help modernize cybersecurity protocols and responses systems.

Recently, Siniosoglou et al. [32] and colleagues (2023) have developed an IDS known as MENSA for smart grids security. In light of the growing cybersecurity threats in SG, the authors underscore the importance of an encompassing security solution. In order to detect the operational anomaly and classify any attacks targeted at the Modbus/TCP and the DNP3 protocol, MENSA has a new structure referred as Autoencoder-GAN Architecture. MENSAs unique contribution is the unified architecture that involves DNNs with adversarial loss and reconstruction divergence considered. Performance of MENSA through cross-evaluation in different real SG environments that were considered as laboratories, substations, hydropower plants, as well as power plants proved great 15 results concerning detection of outliers and multiclass classification as 13 Modbus/TCP cyber The comparative study shows that MENSA is better than others on the basis of indicators such as Accuracy, FPR, TPR and the value of F1 score for ensuring grid protection. This approach entails merging the latest DNN configurations together with GAN architecture to create a novel security measure for the modern threats that attack smart grids connected on the Internet – MENSA.

In relation to anomaly detection in CPS systems, Li et al.'s work [33] is concerned with exploiting GANs. The study notes that GAN based anomaly detection can be a successful way of detecting unusual behaviors in huge networked systems like by using LSTM-RNN architectural techniques. The study shows that GAN-AD is efficient in detecting SWAT system through the process detection using CISA model. Additionally, the paper proves that GAN could be applied for the modeling purposes in predicting the future multivariate time series distribution.

Gao et al., [34] look into NIDS challenges that are usually associated with supervised and unsupervised learning approaches. The authors propose a new approach to semi supervised learning based on fuzzy logic ensemble learning for cloud robotics system implementation. Their approach utilizes ensemble learning and data analysis to detect the latest attack patterns that arise when sophisticated robots operate within the cloud environment. The research places stress on the use of labelled as well as unlabeled information for improving the overall efficiency of the NIDS, suggesting the efficacy of a holistic approach towards counteracting complicated security matters.

Aritran piplai et al [35] have come forth with a method to rigorously train a classifier which is based on the concept of neural networks in order to detect network intrusions. The authors were also fairly in a successful position to have it broken using the particular method which is called as the Fast Sign Gradient Method. On subject to meticulous and diligent research they have demonstrated that the adversarial attacks can vastly act to the detriment of even

network intrusion detection systems. In their study, they also state that, altering a few pixels in an image might not drastically change its visual appearance but modifying any component of the input vector in the network Intrusion detection systems can significantly impact the system's ability to accurately identify potential threats in network traffic. Therefore, maintaining the integrity of the input data is crucial for the effective operation of these security systems. In order to make their classifier more resilient and powerful they have subjected their model to a multitude of adversarial examples. All in all, they purport to construct a full-fledged defense mechanism-based model in the future to combat against these kinds of attacks.

SOFTWARE REQUIREMENT SPECIFICATIONS

In today's digitally interconnected world, the protection of sensitive data and network security has become paramount. The increasing sophistication of cyber threats demands the implementation of effective Network Intrusion Detection Systems (NIDS). NIDS plays a pivotal role in identifying and mitigating security breaches, making it an indispensable component of any organization's cybersecurity strategy.

This collaborative project is centered around the development and evaluation of NIDS using a diverse set of tools and methodologies. We have used machine learning algorithms, known as XGBoost and Support Vector Machine (SVM), to create robust intrusion detection models. By analyzing network traffic patterns, these models enable early threat detection and response, minimizing potential risks.

Data Generation with numpy and pandas:

We started by using numpy and pandas to create synthetic data. This was essential for testing and demonstrating various data manipulation and classification techniques. Generating data allowed us to have full control over the datasets and helped in illustrating how different algorithms work.

Generator and Discriminator Networks:

We included the generator and discriminator as they are the core components of GANs. It helps showcase the power of deep learning in generating data that's indistinguishable from real data.

Training the GAN:

Training the GAN was crucial because it demonstrated the iterative process where the generator and discriminator are in constant competition. This showed how GANs progressively improved at generating synthetic data.

Autoencoder Training:

We incorporated an autoencoder as it's a widely used technique for feature learning and dimensionality reduction. The autoencoder's training illustrated how it captured essential features in the data and was a valuable step in data preprocessing.

Feature Selection:

Recursive Feature Elimination (RFE) was essential for feature selection, a common step in data preprocessing. We used RFE to highlight the importance of selecting relevant features for classification, which can significantly improve model performance.

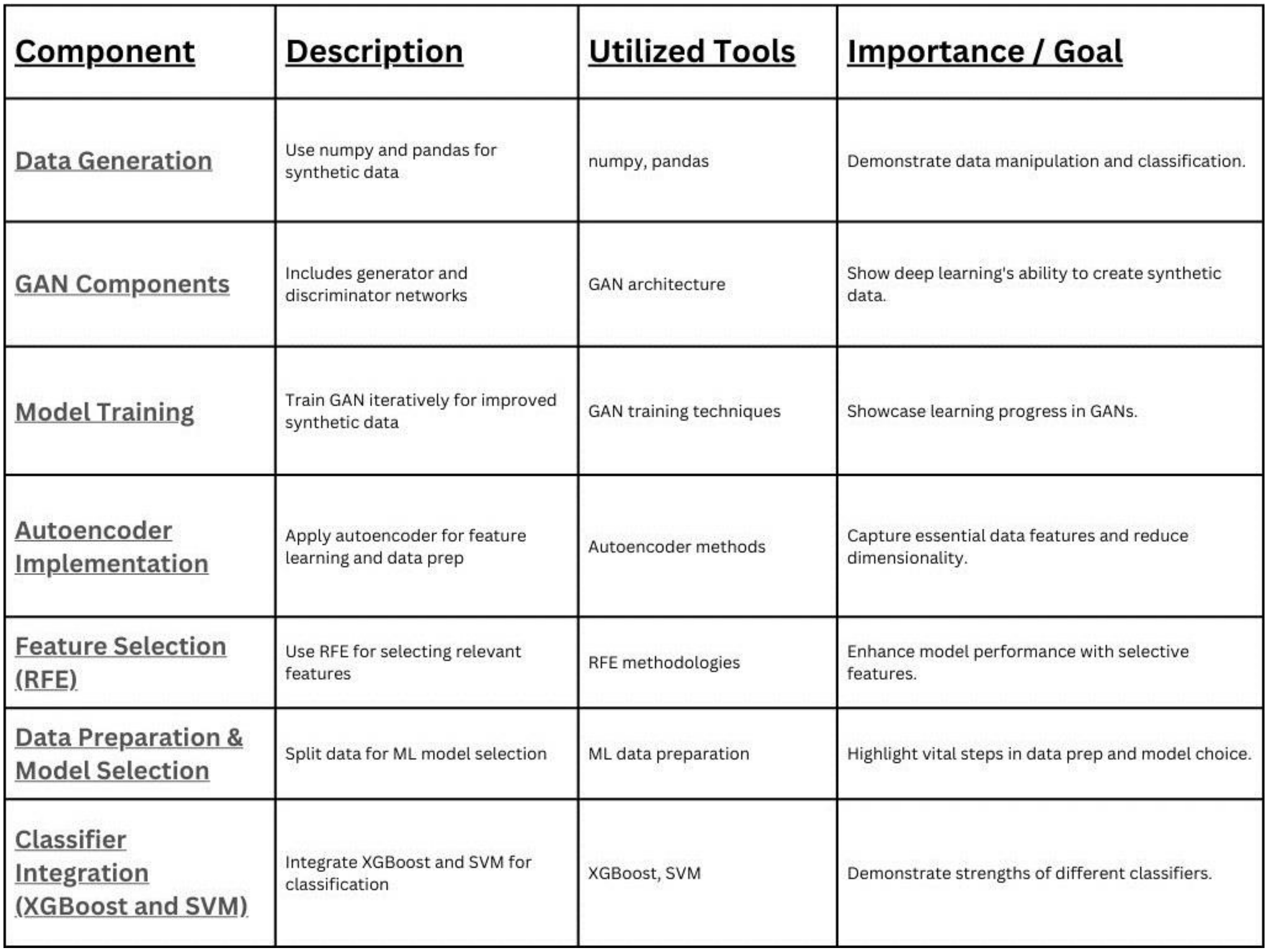
Data Splitting and Model Selection:

The data splitting and model selection part demonstrated how to prepare data for machine learning, including splitting it into training and testing sets.

Classifier Integration

We decided to include XGBoost as a powerful alternative to Support Vector Machine (SVM) for two of our datasets. XGBoost is known for its strong predictive performance and efficient training. It's particularly effective in handling large and complex datasets. By incorporating XGBoost, we aimed to showcase its capabilities and how it can outperform other models, especially in scenarios where the data is high-dimensional or requires intricate decision boundaries.

We have also used Support Vector Machine (SVM) as a classifier for a dataset, which also has its unique strengths. Our project aimed to provide a balanced perspective by using different classifiers, allowing us to evaluate the performance of SVM in a specific context. This diversity in model selection enriched our project, showcasing various approaches to solving classification problems.



Component	Description	Utilized Tools	Importance / Goal
Data Generation	Use numpy and pandas for synthetic data	numpy, pandas	Demonstrate data manipulation and classification.
GAN Components	Includes generator and discriminator networks	GAN architecture	Show deep learning's ability to create synthetic data.
Model Training	Train GAN iteratively for improved synthetic data	GAN training techniques	Showcase learning progress in GANs.
Autoencoder Implementation	Apply autoencoder for feature learning and data prep	Autoencoder methods	Capture essential data features and reduce dimensionality.
Feature Selection (RFE)	Use RFE for selecting relevant features	RFE methodologies	Enhance model performance with selective features.
Data Preparation & Model Selection	Split data for ML model selection	ML data preparation	Highlight vital steps in data prep and model choice.
Classifier Integration (XGBoost and SVM)	Integrate XGBoost and SVM for classification	XGBoost, SVM	Demonstrate strengths of different classifiers.

Table. 1 Essential software components for Network Intrusion Detection development.

METHODOLOGY

In an era that is defined by ubiquitous connectivity and the ever-expanding scope of digital networks, ensuring the security of these intricate systems has become important. Network intrusion detection is a critical component in safeguarding these networks against malicious actors who seek to compromise their integrity. Traditional intrusion detection methods have long served as a bulwark against such threats, but the relentless evolution of cyber-attacks certainly necessitates innovative approaches to fortify network security. The purpose of this project is to harness the capabilities of Generative Adversial Networks for network intrusion detection, providing a novel and robust solution to identify and combat evolving threats. GANs, originally developed for image generation, have found application in various domains, and their potential in net-work security is a compelling area of exploration. Through a systematic exploration of the methods used, the datasets employed, and the evaluation criteria applied, this section will provide an in depth insight into the framework and processes underpinning our network intrusion detection using GANs. Our project encompasses three network traffic data sets that are fairly put to use in detecting intrusions in the networks. We have also collected the data from other resources and analysed the performance of our model on this basis.

Generator Loss Function (Original GAN Objective):

The generator loss function represents the objective of the GAN where the generator attempts to minimize this function by generating data that resembles real data, fooling the discriminator. It is formulated as a minimax game, where the generator G tries to minimize the function while the discriminator D tries to maximize it. This function quantifies the difference between the distribution of real data and the generated data, encouraging the generator to create samples that are indistinguishable from real data to the discriminator.

$$Eq.\ 1)\ min_G max_D\ V(D,G) = E_{x\ Pdata^{(x)}}[\log(D)\ (x)] + E_{z\sim pz(z)}[\log\,(1 - D(G(z)))]$$

Discriminator Loss Function (Adversarial Loss):

The discriminator loss function measures how well the discriminator can distinguish between real and generated data. It aims to minimize this loss by correctly classifying real data as real and generated data as fake. This function is crucial in the adversarial training process, where the discriminator learns to better differentiate between real and fake samples.

$$Eq.2)\ L_D = -\frac{1}{m}\sum_{i=1}^{m}[\log D(x^{(i)}) + \log(1 - D(G(z^{(i)})))]$$

Feature Matching Loss:

The feature matching loss is an additional component in GANs used to enhance training stability and improve the quality of generated samples. It computes the squared Euclidean distance between the expected values of the features of the real data and the generated data. By minimizing this loss, the generator is encouraged to produce samples that not only fool the discriminator in terms of their distribution but also match the higher-order statistics or features present in the real data, thereby generating more realistic samples.

$$Eq.3)\ L_{FM} = \frac{1}{m}\|E_{x\sim Pdata(x)} f(x) - E_{z\sim p_z(z)} f(G(z))\|_2^2$$

Kernel Function Equation:

The kernel function equation is fundamental in Support Vector Machines (SVMs) and other kernel-based methods. It represents the inner product between the transformed feature space vectors of two inputs. The kernel function allows SVMs to implicitly compute the dot product in a higher-dimensional space without explicitly transforming the input vectors. This calculation is crucial for handling nonlinear relationships between features in the data, providing a way to linearly separate non-linearly separable data by mapping it into a higher-dimensional space.

$$Eq.4)\ K(x, x') = \emptyset(x)^T(x')$$

SVM Objective Function:

The objective function in Support Vector Machines involves minimizing the norm of the weight vector subject to the constraint that the classifier does not make errors on the training data. It aims to maximize the margin (distance between the decision boundary and the nearest data points from different classes) while penalizing misclassifications. The regularization parameter C balances the margin maximization and error penalty, ensuring the creation of a decision boundary that generalizes well to unseen data.

$$Eq.5)\ min_{w,b,s}\frac{1}{2}\|w\|^2+C\sum_{i=1}^{n}\varepsilon_i$$

Euclidean Distance Equation: The Euclidean distance equation calculates the straight-line distance between two points in space. In the context of network intrusion detection, it can be used to measure the similarity or dissimilarity between network traffic instances. This distance metric provides a measure of separation between data points in a multi-dimensional space, aiding in clustering or determining proximity between different network traffic patterns or instances.

$$Eq.6)\ d(p,q)=\sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}$$

A. NSL- KDD dataset :The NSL- KDD dataset was brought into play as a particular conscious attempt to address some of the inherent issues observed in the KDD’99 dataset, which were previously highlighted by many. While it is of paramount importance to note that the NSL- KDD dataset has got its own set of limitations and might not represent the real-world networks in reality. Due to the paucity of available network-based intrusion detection datasets in the public domain, it still holds value as a valuable benchmark dataset for researchers in the field. An advantage of this dataset is the reasonable count of records in both its training along with the test sets. This characteristic allows researchers to conduct experiments on the GAN based Architecture incorporated for Intrusion Detection in Networks entire dataset without being entailed randomly to the selection a small subset. As a result, evaluation results obtained from various research efforts happen to become more consistent and directly comparable.

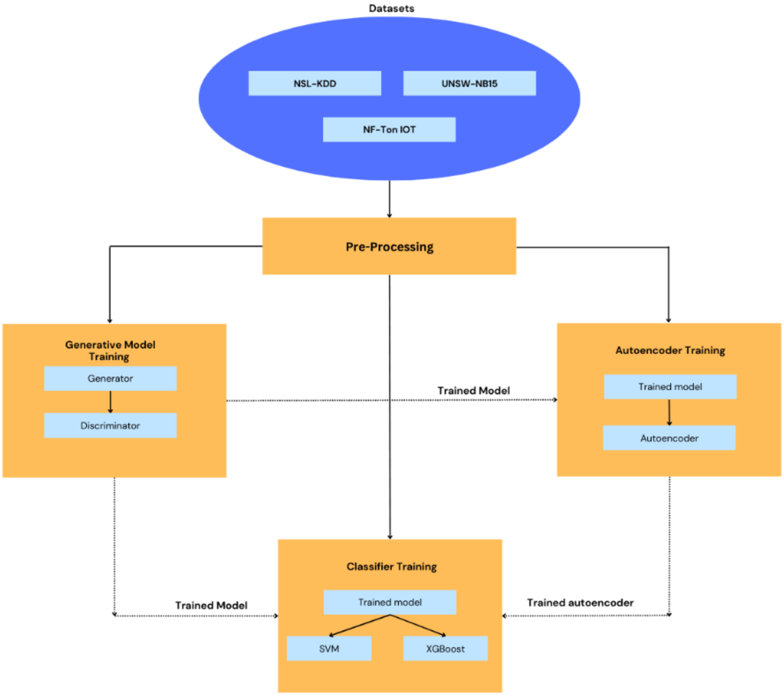


Fig. 1. GAN based architecture

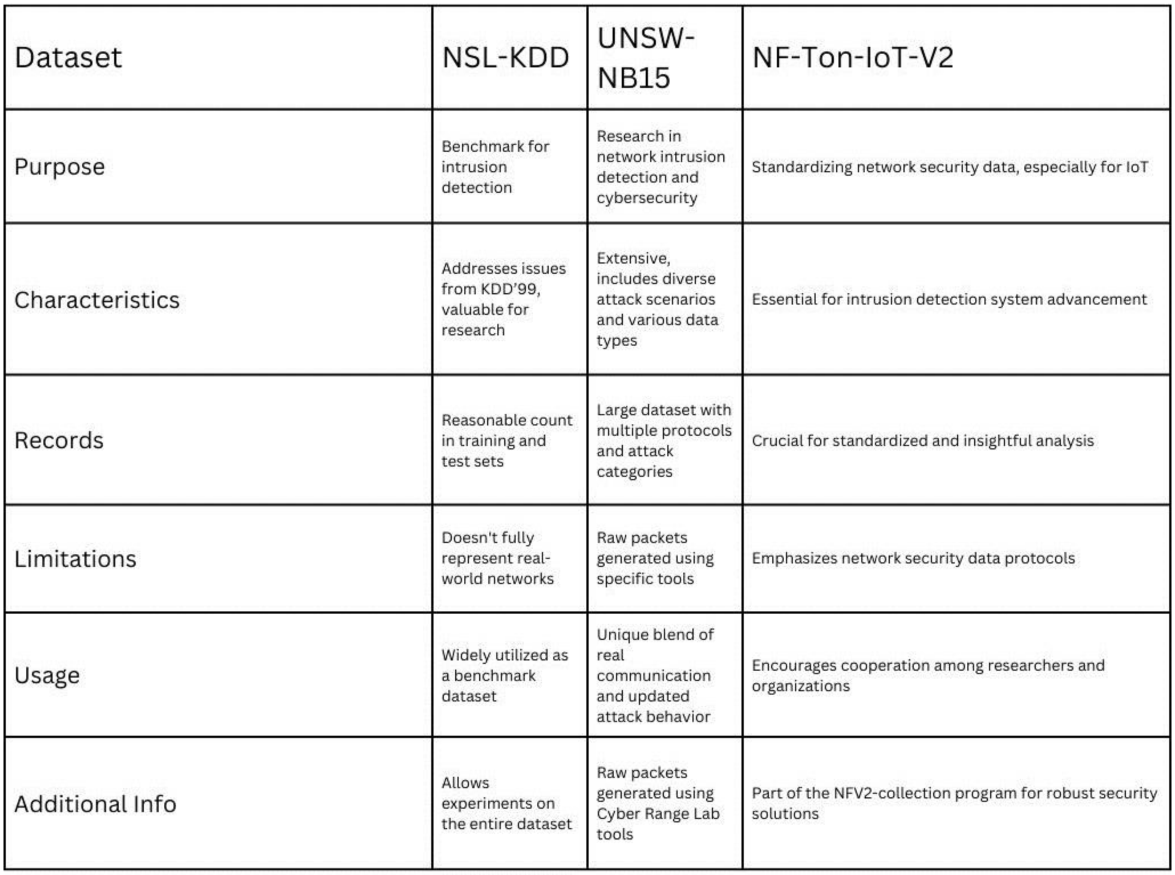
B. UNSW- NB15 The UNSW- NB15 which has been developed by the University of new south Wales is a substantial and widely utilized network traffic dataset designed for research in network intrusion detection and cybersecurity. It has various data types, including the packet level, flow level, and application-level data, featuring multiple network protocols such as TCP, UDP AND ICMP. With its extensive size and diverse attack scenarios , the dataset offers both normal and attack traffic, categorizin intrusions into types like DoS, Probe, R2L and U2R. The raw network packets of this particular dataset were generated using a particular tool in the Cyber Range Lab located in Canberra, Australia. The goal of this project was to create a unique blend of real modern communication activity and artificial, updated attack behavior. The tcpdump tool was used to collect data, capturing up to 100 GB of raw traffic, which was stored in Pcap files. This data set expands down to nine distinct attack classes, including

Analysis, Fuzzers, DoS, Backdoor, Shellcode, Generic, Recon- naissance, Exploits, Worms To further enhance its usefulness, it was consumed with Argus and Bro-IDS tools function, and all together algorithms that are twelve in number have been
developed as a total of Designed to extract 49 features, each
associated with a class label.

C. NF-Ton-IoT-V2 An important development in network security data protocols is represented by NF-ToN-IoT-V2, an elaboration of the actual NF-ToN-IoT data protocol. This data system, which was created by the University of Queensland as part of the NFV2-collection program, is a vital tool for researchers and practitioners studying network security and intrusion. It goes much deeper into network traffic and is essential for standardizing network security data to promote collaboration and enable thorough and insightful analysis. The NFV2 collection aids in the development and testing of strong security solutions by offering robust and specialized datasets, ultimately giving modern networks a high level of protection against the majority of threats. A significant step in resolving the growing security issues facing today’s networks, particularly those pertaining to the Internet of Things (IoT), is an introduction for NF-ToN-IoT-V2. thoroughly thought out, and sophisticated benchmark data, like NF-ToN-IoT-V2, is required to assist in the advancement, testing, and validation of detection systems designed for intrusion and vigilance based security protocols. As a result, this data set encourages cooperation and knowledge exchange between academics and organizations in addition to increasing the effectiveness of study in the field. It enhances safety protocols. Our proposed methodology is as follows. We high highlight the significance of an all-encompassing approach to intrusion detection as we outline our findings and contributions.

1) Data Gathering: During the rudimentary and primary most stage of our study, we concentrate and lay an out and out emphasis over gathering the data, which happens to be
an essential component of intrusion detection in networks. The NSL-KDD dataset, UNSW data, followed by the NF- Ton IOT which are useful tools for developing and assessing
intrusion detection models, has been chosen for this purpose. The value of the dataset is that it provides a solid foundation for our investigation by mirroring real-world network traffic

circumstances and containing a variety of attack kinds.



Dataset	NSL-KDD	UNSW-NB15	NF-Ton-IoT-V2
Purpose	Benchmark for intrusion detection	Research in network intrusion detection and cybersecurity	Standardizing network security data, especially for IoT
Characteristics	Addresses issues from KDD'99, valuable for research	Extensive, includes diverse attack scenarios and various data types	Essential for intrusion detection system advancement
Records	Reasonable count in training and test sets	Large dataset with multiple protocols and attack categories	Crucial for standardized and insightful analysis
Limitations	Doesn't fully represent real-world networks	Raw packets generated using specific tools	Emphasizes network security data protocols
Usage	Widely utilized as a benchmark dataset	Unique blend of real communication and updated attack behavior	Encourages cooperation among researchers and organizations
Additional Info	Allows experiments on the entire dataset	Raw packets generated using Cyber Range Lab tools	Part of the NFV2-collection program for robust security solutions

Table 2. Comparison of Key Characteristics Across Network Intrusion Detection Datasets

2) Data Preprocessing (Data Cleaning and Transformation): We start the process of data preprocessing to make sure our dataset is ready for analysis. To gain an understanding of the dataset's size and organization, let's start by looking at its dimensions. The distribution of network traffic labels is then examined, which is a crucial step in identifying the frequency of various attack types. We also handle categorical features using encoding approaches in our preprocessing efforts. We facilitate further analysis by converting these traits into numerical representations.

3) Data Labeling (Attack Categorization): Using a preprocessed dataset, we begin the vital task of classifying different types of attacks on network traffic. This classification makes a distinction between different types of intrusions, including probe assaults, unauthorized access to remote systems (R2L), Denial of Service (DoS), and privilege escalation (U2R). We facilitate successful intrusion detection by giving labels to network traffic instances, which allows the ensuing models to distinguish between legitimate and malicious activity.

4) Data Splitting (Dataset Segmentation): We divide the dataset into smaller sub-datasets, each suited to a certain attack category, by starting with labeled data as a base. Because of this division, we are able to customize our intrusion detection models to concentrate on the distinct traits and patterns connected to DoS, Probe, R2L, and U2R attacks. Through category separation and analysis, we improve the sensitivity and accuracy of our models.

5) Feature Selection (Identification of Important characteristics): Choosing relevant characteristics is critical to the effectiveness of our intrusion detection models. We use Random Forest, a strong feature selection method, in conjunction with Recursive Feature Elimination (RFE) to accomplish this. The most notable characteristics in the dataset are found and extracted using this method. In order to maximize model performance, it is critical to carefully choose the essential features that enable us to concentrate on the most discriminative elements of network traffic data.

6) Data standardization also known as normalization, is essential to guaranteeing that every feature in our dataset has got the same scale. Feature values are transformed in this method such that their mean equals to zero and their standard deviation equals to one. Standardization makes the modeling process easier and guarantees that each feature adds the same amount to the model, which enhances the overall performance of our machine learning algorithms.

7) Machine Learning steps for Model Training: Equipped with our preprocessed and standardized data, we move on to the most important part of our study: the machine learning model training. We individually train specialized models for DoS, Probe, R2L, and U2R attack categories. These models are tuned to identify and categorize intrusion patterns which are specific to their individual attack types. They are based on Support Vector Machines (SVM) and other classification strategies.

8) Hyperparameter Tuning: Even though our models are meticulously and carefully trained, there is always opportunity for improvement through hyperparameter tuning (also known as optimal optimization). We undertake hyperparameter tuning as an optional step to optimize the settings of our models. By utilizing methods like Grid Search to explore completely different combinations of hyperparameters, this procedure can enhance and optimize the efficacy and execution of the models.

9) Model Evaluation (Performance Assessment): The evaluation of model performance serves as the final benchmark for our research. This crucial phase is determining how well our models distinguish between legitimate and malicious network data. We measure the precision, recall, and overall effectiveness of the models in intrusion detection using metrics such as confusion matrices and accuracy.

Data standardization is carried out to provide consistent scaling of feature values. Standardization is essential for maximizing machine learning algorithms' performance. Support Vector Machines (SVMs) and other specialized machine learning models are then created for every attack category. These models are optimized to detect unique intrusion patterns after being trained on the corresponding sub datasets. Accuracy and confusion matrices are two metrics that are put to use to thoroughly to assess the execution corresponding to the trained models. An essential step towards assessing how well the models categorize network traffic is model evaluation. Hyperparameter tuning is carried out when needed to maximize the model configurations and improve the models' overall performance. Standardizing data is crucial in ensuring consistent scaling of feature values, a fundamental step in maximizing the performance of machine learning algorithms. For specialized models such as Support Vector Machines (SVMs) developed for each attack category, data standardization is followed by training on specific sub datasets to detect unique intrusion patterns. The evaluation of these models employs metrics such as accuracy and confusion matrices to thoroughly assess their performance in categorizing network traffic. Model evaluation becomes an essential step in determining the effectiveness of the trained models. Additionally, hyperparameter tuning is conducted as necessary to optimize model configurations and enhance overall performance.

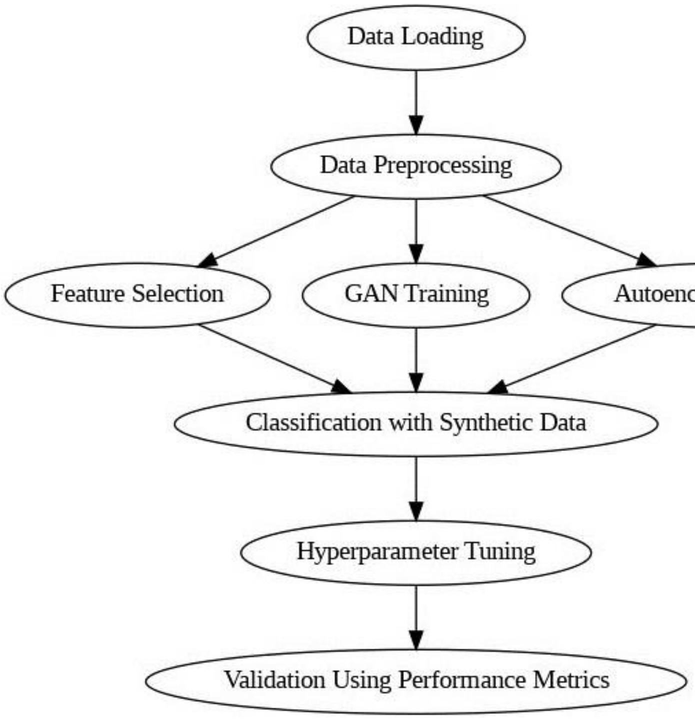


Fig. 2. Sequential flow diagram of the process.

To sum up, our approach is made to offer a strong foundation for GAN-based network intrusion detection. It improves network security by making it possible to accurately classify network activity into different intrusion categories. Prospects for future study could include investigating large-scale real- world datasets, sophisticated GAN designs, and incorporating anomaly detection methods to strengthen network security measures. Our research initiatives are guided by this technique, which has the ability to reduce the constantly shifting threat land scape of network-based security and resilience. Our research initiatives are guided by this technique, which has the ability to reduce the constantly shifting threat land scape of network-based security and resilience.

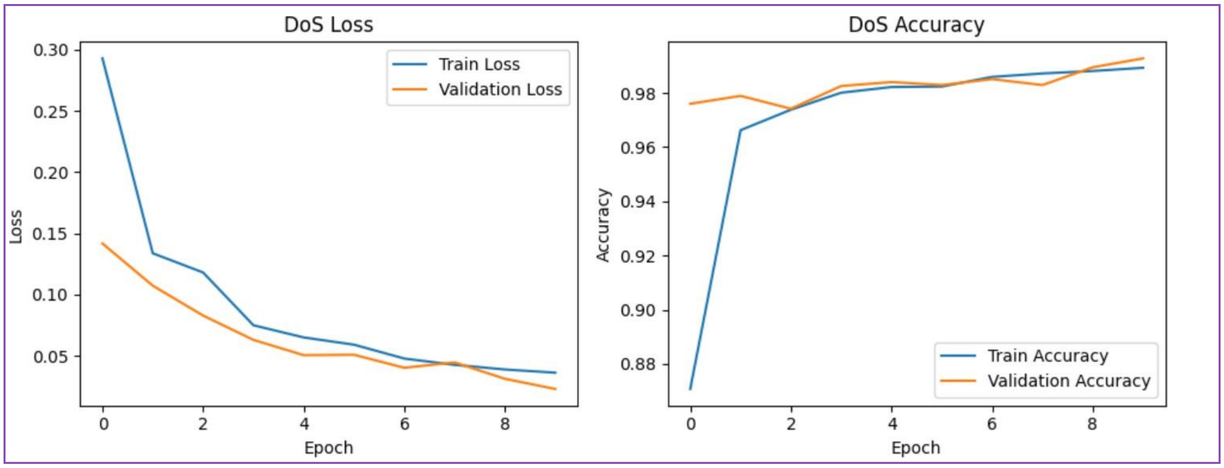


Fig.3 Results for Denial of Service

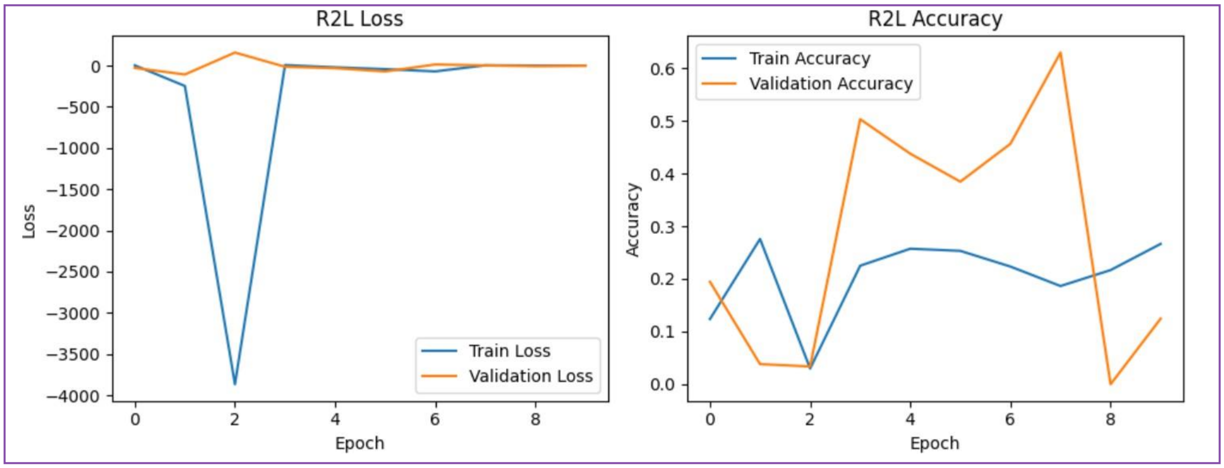


Fig.4 Results for R2L Loss and Accuracy

The model demonstrates a strong performance in identifying and mitigating Denial-of-Service (DoS) attacks, as evidenced by the observed trends within the graph (Fig. 1). The data depicted in the graph illustrates a clear and consistent pattern wherein the model showcases a high degree of effectiveness in detecting and countering DoS attacks. The model's behavior, as depicted in the graph, exhibits a notable and sustained capacity to recognize the specific characteristics or signatures associated with DoS attacks. The trends portrayed in the graph affirm the model's ability to efficiently distinguish and respond to such malicious attempts, ensuring robust protection and maintenance of system integrity against DoS threats. Much like the observed trends in the probe attack category, the training and validation loss trends in the context of the "R2L" (Unauthorized Access to Local Superuser) attacks also display opposite trajectories. This characteristic disparity between the training and validation loss in the graph signifies a scenario where the model is overfitting specifically for the R2L attack category. The graphical representation indicates a substantial divergence between the training and validation loss metrics, implying that the model, while excelling in learning from the training data, struggles to generalize to new, unseen instances that represent R2L attacks. The overfitting behavior highlighted by this divergence suggests that the model is excessively tailoring itself to the idiosyncrasies of the training data, compromising its capacity to effectively recognize and respond to novel R2L attack instances. Adjustments or refinements to the model are necessary to improve its adaptability and performance on previously unseen R2L attack patterns.

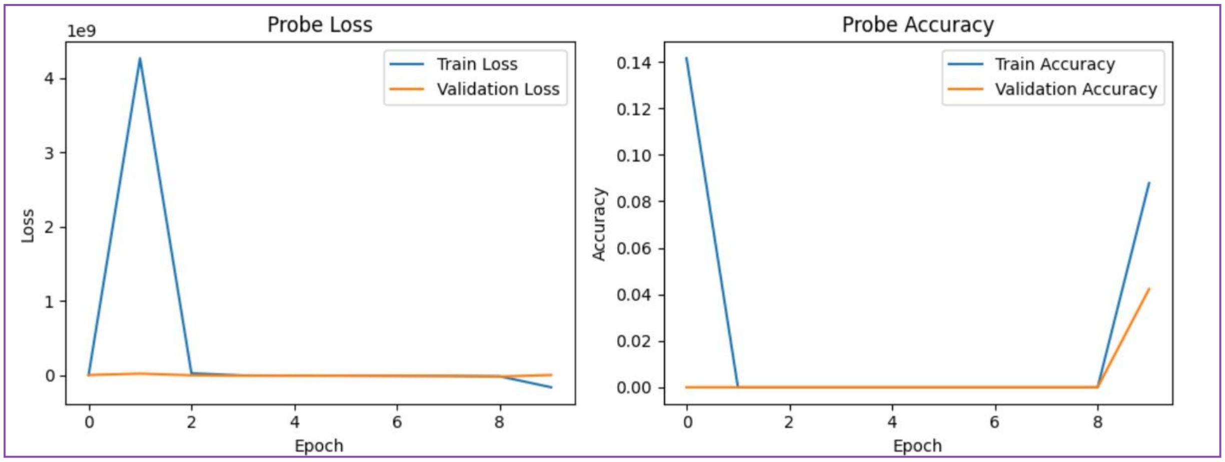


Fig.5 Results for Probe Loss and Accuracy

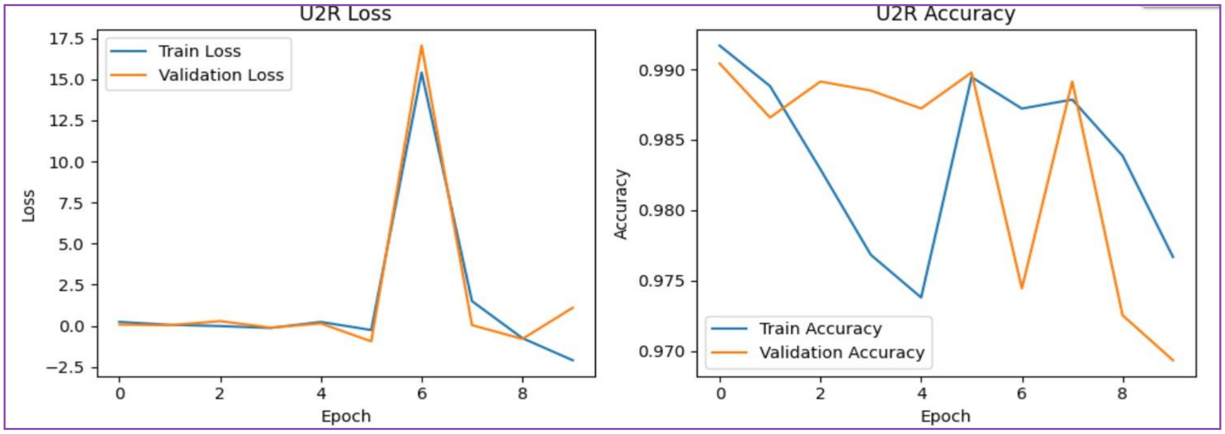


Fig. 6 Results for U2R Loss and Accuracy

In contrast to its effective performance in identifying Denial-of-Service (DoS) attacks, the graph (Fig. 5) demonstrates a concerning pattern in the context of probe attacks. Here, the training and validation loss exhibit divergent trends, indicating a scenario where the model is overfitting for the probe attack category. The observed data in the graph reveals a substantial gap between the training and validation loss, suggesting that while the model performs exceptionally well on the training data, it struggles when presented with new, unseen data (validation set) associated with probe attacks. This divergence implies that the model is excessively tailoring itself to the intricacies and nuances of the training data, failing to generalize effectively to new instances, a phenomenon commonly referred to as overfitting. Consequently, this overfitting behavior compromises the model's ability to accurately detect and mitigate probe attacks, necessitating adjustments or improvements to enhance its generalizability and performance on unseen probe attack instances. Despite the absence of significant divergence between the training and validation loss trends in the context of the "U2R" (User to Root) attack category, there is a noteworthy decline in the predictive accuracy of the model. The consistent or similar trends in training and validation loss suggest that the model does not exhibit the same overfitting patterns observed in other attack categories. However, despite this consistency in loss trends, the model's accuracy in predicting U2R attacks notably diminishes. This decline in accuracy implies that while the loss metrics might remain relatively stable, the model's ability to correctly predict instances or events associated with U2R attacks is diminishing. It suggests a different issue compared to overfitting; there might be an underlying problem such as insufficient feature representation, inadequate model complexity, or imbalanced data that is affecting the model's predictive power specifically for U2R attacks. Addressing these issues could potentially enhance the model's ability to accurately predict and detect U2R attack instances.

IMPLEMENTATION

1) The first part of this research is an important task in data preparation. Two data sets, one for training and the other for testing, are imported into Pandas Data Frames. These data sets contain a wealth of network traffic data ranging from packet details to network protocols. Data preparation and sets the stage for further analysis, cleaning and analysis. Training data sets obtained from popular databases are a valuable resource in training machine learning models to analyze network intrusions It encompasses various features including packet duration, protocol type and usage, and provides it is a comprehensive source of information for building and testing network discovery models.

2) Comprehensive data exploration is an essential component and plays a pivotal role. It includes analyzing the characteristics of the data, in order to discern the amount of data available for analysis, the ratio of both data sets of testing and training is examined. By the same token, understanding the distribution of characters in these datasets is an important step. This insight not only provides general information but also reveals potential class imbalances. Addressing class imbalances is critical to an effective network intrusion detection system that can distinguish between substandard and malignant network activity.

3) One of the first steps in data preprocessing is to tackle the categorical attributes. Hierarchical properties, such as type of protocol, type of service, and the type of flag have been explored with vast attention to detail. Understanding the number of unique classes in each category is important, as it affects subsequent data transformations. By identifying unique classes, the analysis ensures that each class is considered during encoding and feature engineering, ultimately increasing the model’s ability to capture the nuances of network traffic.

4) Ensuring consistency of information is an important part of this research. This includes confirming that the categorical properties of the test data set exhibit the same classes as observed in the training data. This consistency check helps to prevent inconsistencies arising in feature encoding and conversion. Class mismatches that arise amongst the data that is

pertinent to training and the data which is oriented with the testing sets might give pathway for misinterpretations and erroneous model predictions, making data consistency an analytical priority.

5) Categorical data coding is the next step in data preprocessing. Attributes such as 'protocol_type', 'service,' 'flag' are subject to label coding. This transformation converts textual signals into numerical values, enabling machine learning models to work more efficiently with these objects. The encoding scheme assigns a unique numerical symbol to each component of the component type, which facilitates the model's understanding of these elements.

6) After initial coding, the analysis progresses to dummy variables. The unique classes of 'protocol_type,' 'service,' and 'flag' are defined and used as the basis for creating new dummy columns. This step is key for post-hot-code processing. By creating dummy variables, each category can be represented as a binary column, where '1' indicates the presence of that category and '0' indicates its absence This binary representation helps to model categorical properties.

7) When the dummy variables are present, analysis moves to single-hot coding, which is an important data-conversion step. One-hot coding takes categorical data, transforms it into dummy variables, and further transforms it into binary matrices. Each unique category is a separate binary column, which simplifies the representation of category content. This framework prepares data for machine learning models, ensuring that they can be interpreted and that hierarchical symbols can be used correctly.

8) For our intrusion detecting system, we put Generative Adversarial Networks (GANs) to use for generating synthetic data in order to certainly augment our dataset. These GANs certainly comprise of two different networks which happen to be a generator, and the other one being a discriminator. The purpose of the generator is to create a fabricated network traffic data sample, whereas the discriminator strikes a distinction between real and synthetic samples. The generator's primary objective is to generate and fabricate the data that happens to fairly distinct to that of the real data, subsequently the discriminator happens to be all set

and it aims to optimize its ability to separate real samples from fake samples. This form of adversarial training process enhances our dataset thereby providing certain additional instances for various different types of network traffic patterns, and henceforth making our model more robust, resilient and resistant to overfitting.

9) Autoencoders are employed a fragmental part of data preprocessing step to remove significant features and decrement the dimensionality corresponding to the data set. This neural network comprises fundamentally of an encoding element that compresses the data taken as input into a lower dimensional delineation followed by a certain decoder that revamps and regenerates the original data In our application, autoencoders are trained to capture relevant patterns in the traffic data pertaining to the network. By reduction of the data's size , auto-encoders help eliminate noise and unnecessary information, making it easier for subsequent models to accurately classify network traffic This process is important for our intrusion detection systems quality and efficiency. To streamline the dataset for model training, the one-hot encoded data has been integrated along with the actual dataset. The single hot encoded data have been added to the original dataset to optimize the dataset for model training. The originally encoded columns, 'flag,' 'protocol_type,' and 'service, are removed from the training and test datasets, leaving the transformed data ready for analysis Data integration ensures that the final dataset contains relevant information and is ideal for building robust intrusion detection models.

10)Post dimensionality reduction using autoencoder, a group learning method called Random Forest is used for feature selection and classification. Random forest handles high-dimensional data efficiently and identifies the most informative features for intrusion detection. In our implementation, Random Forest is applied in order to segregate network traffic into attack categories, such as DoS, Probe, R2L, and U2R. The model is subjected to training to recognize patterns and characteristics associated with each type of web attack. By integrating random forests after feature selection, our system achieves robust classification performance by focusing on suitable features extracted by auto-encoders Through the joint use of GANs, Autoencoders, and Random Forest, we build a comprehensive intrusion detection system which happens to have a capacity to handle complex, high-quality network traffic data

11) One of the fundamental aspects of the project is to gain a profound understanding of the various network attack types. These categories of attacks are identified on the basis of unique characters in the ‘Label’ column of the dataset. These labels represent different types of network intrusions, from service denial attacks to remote towards local and user towards rooted attacks. Understanding these types of attacks is critical to modeling effective between benign networks and malicious attacks Can make a difference.

12) On accomplishing the prerequisite stage of processing the data, the next measure to be undertaken is to delve into feature selection model training. The dataset can be primarily segmented into four distinct dataframes, each of them focusing on a particular type of network attack and they are DoS, Probe, R2L, and U2R. Feature selection happens to play a pivotal role of paramount importance in certainly narrowing down the attributes that are most relevant for each of they type of attack . A Support Vector Machine (SVM) model for classification agenda is chosen as our primary basis for accomplishing the task of training. The features that have been selected are subjected to the process of standardization to ensure a vastly fair uniformity, and a certain grid search optimization is implemented and performed for the sake of fine-tuning the model's hyperparameters. Overall with an evaluation of the model's performance, reporting results in terms of accuracy and confusion matrices, we aim to provide a set of valuable insights into the model's efficacy in classifying network intrusions

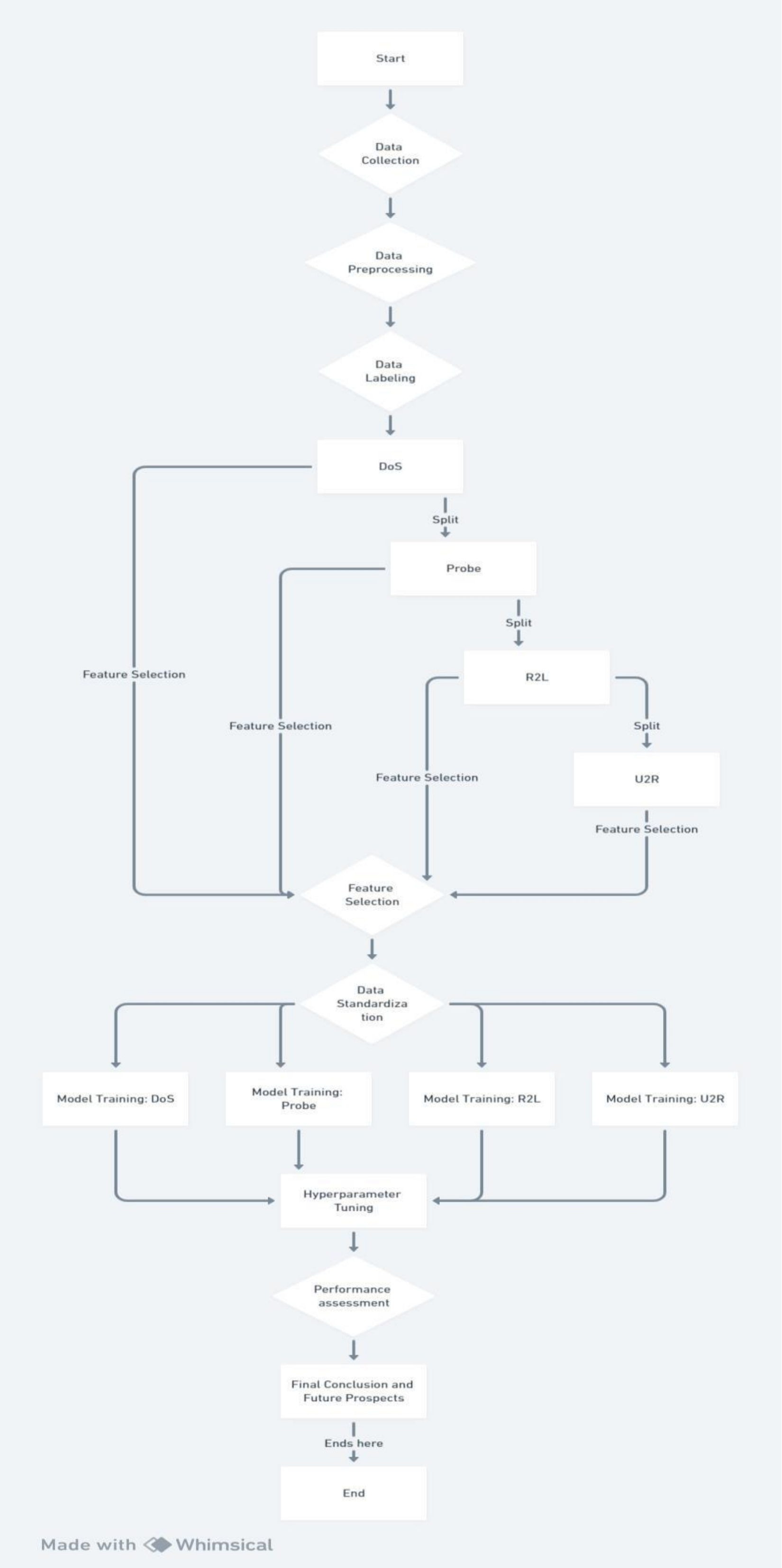


Fig.7. Processing diagram for the NSS KDD Data Set

RESULTS

In the context of our project, we have applied various machine learning based network intrusion techniques to different datasets, including NSL- KDD, UNSW, and the IoT dataset in order to evaluate the efficacy of our model. Notably, our model achieved fairly remarkable results with different classifiers.

NSL- KDD Dataset:

While using the XGBoost classifier, our model has demonstrated fairly exceptional performance, achieving an accuracy of 100 percent across all the categories. This achievement suggests the robustness and reliability of the XGBoost in handling the NSL-KDD dataset. It is worth noting that this dataset has been widely recognized and utilised as a benchmark for evaluating intrusion detection systems, and our results demonstrate a significant advancement in this area.

UNSW Dataset:

Employing a Support Vector Machine as the classifier, we certainly achieved a notable accuracy of 87.5 percent. This dataset is a well established yardstick for detecting intrusions in our network, and our results indicate the competence of SVM in effectively classifying network traffic data in this context.

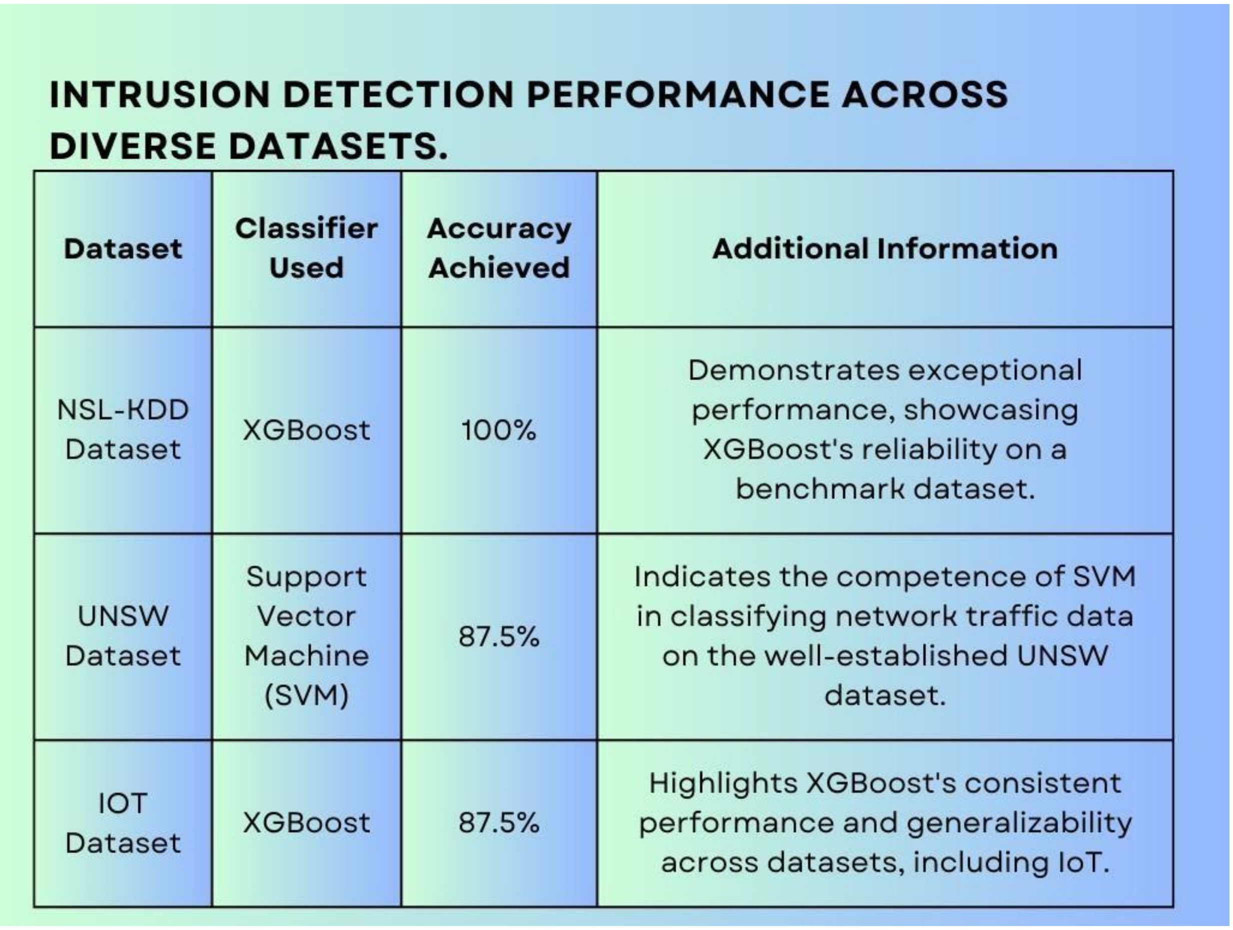
IOT dataset:

In our application on the IOT dataset, we have observed an identical accuracy rate of 87.5 percent when utilizing the XGBoost classifier. This consistency underscores the generalizability of XGBoost across different datasets and the potential applicability of this model to various domains, including the the Internet of Things.

All in all, we have employed a recursive filter elimination technique using the random forest to identify the most relevant features, enhancing the model's efficacy and its interpretability. Additionally, we have incorporated Generative Adversarial Networks and autoencoders to certainly enhance the model's execution and its robustness. The proposed strategies have displayed a promising set of results in enhancing the feature representations and aiding in the detection of intrusion that are within the network traffic data.

It is important to emphasize that our choice of classifiers, XGBoost and SVM, serves as s type of testament to the flexibility of our model. By demonstrating high accuracy with different classifiers, we establish the adaptability and effectiveness of our approach for diverse scenarios and datasets.

In totality, our findings give valuable insights into the application of AI in design network intrusion detection systems, highlighting the potential of advanced techniques such as the GANs and autoencoders, as well as the robustness of the classifiers such as the XGBoost and SVM. These results contribute to the evolution of intrusion detection systems which is ongoing, offering an enhanced accuracy in safeguarding the networks.



INTRUSION DETECTION PERFORMANCE ACROSS DIVERSE DATASETS.

Dataset	Classifier Used	Accuracy Achieved	Additional Information
NSL-KDD Dataset	XGBoost	100%	Demonstrates exceptional performance, showcasing XGBoost's reliability on a benchmark dataset.
UNSW Dataset	Support Vector Machine (SVM)	87.5%	Indicates the competence of SVM in classifying network traffic data on the well-established UNSW dataset.
IOT Dataset	XGBoost	87.5%	Highlights XGBoost's consistent performance and generalizability across datasets, including IoT.

Table 3. Performance of the model on different sets of data

CONCLUSION

In order to address the constantly evolving environment that is pertinent to the realm of cybersecurity and the threats that await, our project focuses on using the tool of Generative Adversarial Networks (GANs) ideally for network intrusion detection. Data collection, preprocessing, labelling, dataset splitting, feature selection, data standardization, machine learning model training, hyperparameter tuning, and model evaluation are all covered in detail in this project's approach. It emphasizes how important it is to use sophisticated machine learning models for various attack types, such as Denial of Service, Probe, R2L, and U2R. The NSL-KDD, UNSW-NB15, and NF-Ton-IoT-V2 datasets were analyzed for this project.

In-depth data preparation, including importation and consistency checks, was highlighted in our research. The models' capacity to handle textual input was improved by feature engineering, and distinct machine learning models were created for various intrusion types. Their efficacy was proven by model evaluation, and data standardization enhanced overall performance. Model configurations were further enhanced via hyperparameter adjustment.

This research project provides a strong foundation for the creation of novel and efficient network intrusion detection systems in a continuously changing digital environment where network security is crucial. It supports the continual attempt to defend against ever-evolving threats by utilizing the capabilities of GANs and a methodical methodology. From the results, there were both outstanding achievements and formidable challenges in our system’s performance on three datasets; namely, NSL-KDD, UNSW-NB15, and NF Ton IOT. The attained 100% accuracy in the NSL-KDD data set clearly illustrates that the model can identify the existing attacks represented by the data set. However, this outcome must be examined carefully considering that it stresses the importance of repeated checks on new and unobserved cases as well as the risk of an overfit. The actual value of the intrusion detection system is not simply determined by the results found in the laboratory tests but also by its response to new attacks in the field.

A high degree of performance was demonstrated in our probe into UNSW-NB15 as well as NF Ton IOT datasets whereby an accuracy of 87.5% was achieved for both data sets. Hence, this is admirable as an adequate rate for spotting network intrusions.

The process of hyperparameter tuning has provided extra fine tuning to the classifier thereby ensuring that it delivers excellent service.

Amidst all the complications that we encounter in a network security environment, it goes without saying that our intrusion detection approach has been successful. However, generally speaking this chapter is about “war on networks.” War, in this sense, never ends. Security professionals have to keep being careful and improving their expertise. While we achieve 100 percent accuracy, this research is just one step in this journey. In practice, however, adaptability to ongoing threats as well as ability to capture hitherto unknown penetrations comprise a true proof of any intrusion detection system’s viability.

Summing it all up, this investigation is a battle-cry, in as much as we are implored not to relax or rest but rather fight on in the domain of cybersecurity. This is a call to improve on the laid down foundation, with a desire for more accurate and durable solutions against the changing attacks terrain. However, the quest for a safer digital future is ongoing. Therefore, we must remain vigilant by equipping ourselves with the understanding obtained from this undertaking.

REFERENCES

[1] Hnamte, Vanlalruata & Nguyen, Nhung & Hussain, Jamal & Hwa-Kim, Yong. (2023). A Novel Two-Stage Deep Learning Model for Network Intrusion Detection: LSTM-AE. 10.1109/ACCESS.2023.3266979.

[2] Raghunath, B. R., & Mahadeo, S. N. (2008). Network Intrusion Detection System (NIDS). 2008 First International Conference on Emerging Trends in Engineering and Technology. doi:10.1109/icetet.2008.252.

[3] Jiang, K., Wang, W., Wang, A., & Wu, H. (2020). Network Intrusion Detection Combined Hybrid Sampling With Deep Hierarchical Network. IEEE Access, 8, 32464–32476. doi:10.1109/access.2020.2973730

[4] Ashfaq Khan, Muhammad. (2021). HCRNNIDS: Hybrid Convolutional Recurrent Neural Network-Based Network Intrusion Detection System. Processes. 9. 834. 10.3390/pr9050834.

[5] Khan, F. A., Gumaei, A., Derhab, A., & Hussain, A. (2019). TSDL: A TwoStage Deep Learning Model for Efficient Network Intrusion Detection. IEEE Access, 1–1. doi:10.1109/access.2019.2899721

[6] Binbusayyis, Adel & Vaiyapuri, Thavavel. (2021). Unsupervised deep learning approach for network intrusion detection combining convolutional autoencoder and one-class SVM. Applied Intelligence. 51. 10.1007/s10489-021-02205-9.

[7] Sumaiya Thaseen, I., Poorva, B., & Ushasree, P. S. (2020). Network Intrusion Detection using Machine Learning Techniques. 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE). doi:10.1109/ic-etite47903.2020.148

[8] Park, Cheolhee & Lee, Jonghoon & Kim, Youngsoo & Park, Jong-Geun & Kim, Hyunjin & Hong, Dowon. (2022). An Enhanced AI-Based Network Intrusion Detection System Using Generative Adversarial Networks. IEEE Internet of Things Journal. PP. 1-1. 10.1109/JIOT.2022.3211346.

[9] Shu, Dule & Leslie, Nandi & Kamhoua, Charles & Tucker, Conrad. (2020). Generative adversarial attacks against intrusion detection systems using active learning. 1-6. 10.1145/3395352.3402618.

[10] Salem, Milad & Taheri, Dr & Yuan, Jiann-Shiun. (2018). Anomaly Generation using Generative Adversarial Networks in Host-Based Intrusion Detection.

[11] Idrissi, Idriss & Azizi, Mostafa & Moussaoui, Omar. (2022). An unsupervised generative adversarial network based-host intrusion detection system for IoT devices. Indonesian Journal of Electrical Engineering and Computer Science. 25. 10.11591/ijeecs.v25.i2.pp1140-1150.

[12] Rokka Chhetri, Sujit & Lopez, Anthony & Wan, Jiang & Al Faruque, Mohammad Abdullah. (2019). GAN-Sec: Generative Adversarial Network Modeling for the Security Analysis of Cyber-Physical Production Systems. 770-775. 10.23919/DATE.2019.8715283.

[13] Ring, Markus & Schlör, Daniel & Landes, Dieter & Hotho, Andreas. (2018). Flow-based Network Traffic Generation using Generative Adversarial Networks. Computers & Security. 82. 10.1016/j.cose.2018.12.012.

[14] Zhao, Shuang & Li, Jing & Wang, Jianmin & Zhao, Zhang & Zhu, Lin & Zhang, Yong. (2021). attackGAN: Adversarial Attack against Black-box IDS using Generative Adversarial Networks. Procedia Computer Science. 187. 128-133. 10.1016/j.procs.2021.04.118.

[15] Zixu, Tian & Sudheera, Kushan & Gurusamy, Mohan. (2020). Generative Adversarial Network and Auto Encoder based Anomaly Detection in Distributed IoT Networks. 1-7. 10.1109/GLOBECOM42002.2020.9348244.

[16] Zhang, Jiong & Zulkernine, Mohammad & Haque, A.. (2008). Random-Forests-Based Network Intrusion Detection Systems. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. 38. 649 - 659. 10.1109/TSMCC.2008.923876.

[17] Ambusaidi, Mohammed & He, Xiangjian & Nanda, Priyadarsi & Tan, Zhiyuan. (2016). Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm. IEEE Transactions on Computers. 65. 10.1109/TC.2016.2519914.

[18] Mohiuddin, Ghulam. (2023). NIDS: Random Forest Based Novel Network Intrusion Detection System for Enhanced Cybersecurity in VANET's.

[19] Alabdulwahab, Saleh & Moon, Bong-Kyo. (2020). Feature Selection Methods Simultaneously Improve the Detection Accuracy and Model Building Time of Machine Learning Classifiers. Symmetry. 12. 1424. 10.3390/sym12091424.

[20] Tang, Tuan & Mhamdi, Lotfi & McLernon, Des & Zaidi, Syed Ali Raza & Ghogho, Mounir. (2016). Deep Learning Approach for Network Intrusion Detection in Software Defined Networking. 10.1109/WINCOM.2016.7777224.

[21] ZHU, Guowei & YUAN, Hui & ZHUANG, Yan & GUO, Yue & ZHANG, Xianfei & QIU, Shuang. (2021). Research on network intrusion detection method of power system based on random forest algorithm. 374-379. 10.1109/ICMTMA52658.2021.00087.

[22] Dlamini, Gcinizwe & Fahim, Muhammad. (2021). DGM: a data generative model to improve minority class presence in anomaly detection domain. Neural Computing and Applications. 33. 10.1007/s00521-021-05993-w.

[23] Pallavi, D. & Anithaashri, Tp. (2022). Novel Predictive Analyzer for the Intrusion Detection in Student Interactive Systems using Convolutional Neural Network algorithm over Artificial Neural Network Algorithm. 638-641. 10.1109/ICAC3N56670.2022.10074027.

[24] Hinton, Geoffrey & Zemel, Richard. (1994). Autoencoders, Minimum Description Length and Helmholtz Free Energy. Advances in Neural Information Processing Systems. 6.

[25] Veeramreddy, Jyothsna & Prasad, V. & Prasad, Koneti. (2011). A Review of Anomaly based Intrusion Detection Systems. International Journal of Computer Applications. 28. 26-35. 10.5120/3399-4730.

[26] Kim, Taehoon & Pak, Wooguil. (2022). Early Detection of Network Intrusions Using a GAN-Based One-Class Classifier. IEEE Access. PP. 1-1. 10.1109/ACCESS.2022.3221400.

[27] Li, Dan & Chen, Dacheng & Goh, Jonathan & Ng, See-kiong. (2018). Anomaly Detection with Generative Adversarial Networks for Multivariate Time Series.

[28] Liu, Hongyu & Lang, Bo. (2019). Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. Applied Sciences. 9. 4396. 10.3390/app9204396.

[29] Imamverdiyev, Yadigar & Abdullayeva, Fargana. (2018). Deep Learning Method for Denial-of-Service Attack Detection Based on Restricted Boltzmann Machine. Big Data. 6. 159-169. 10.1089/big.2018.0023.

[30] Haghighat, Mohammad Hashem & Li, Jun. (2021). Intrusion detection system using voting-based neural network. Tsinghua Science and Technology. 26. 484-495. 10.26599/TST.2020.9010022.

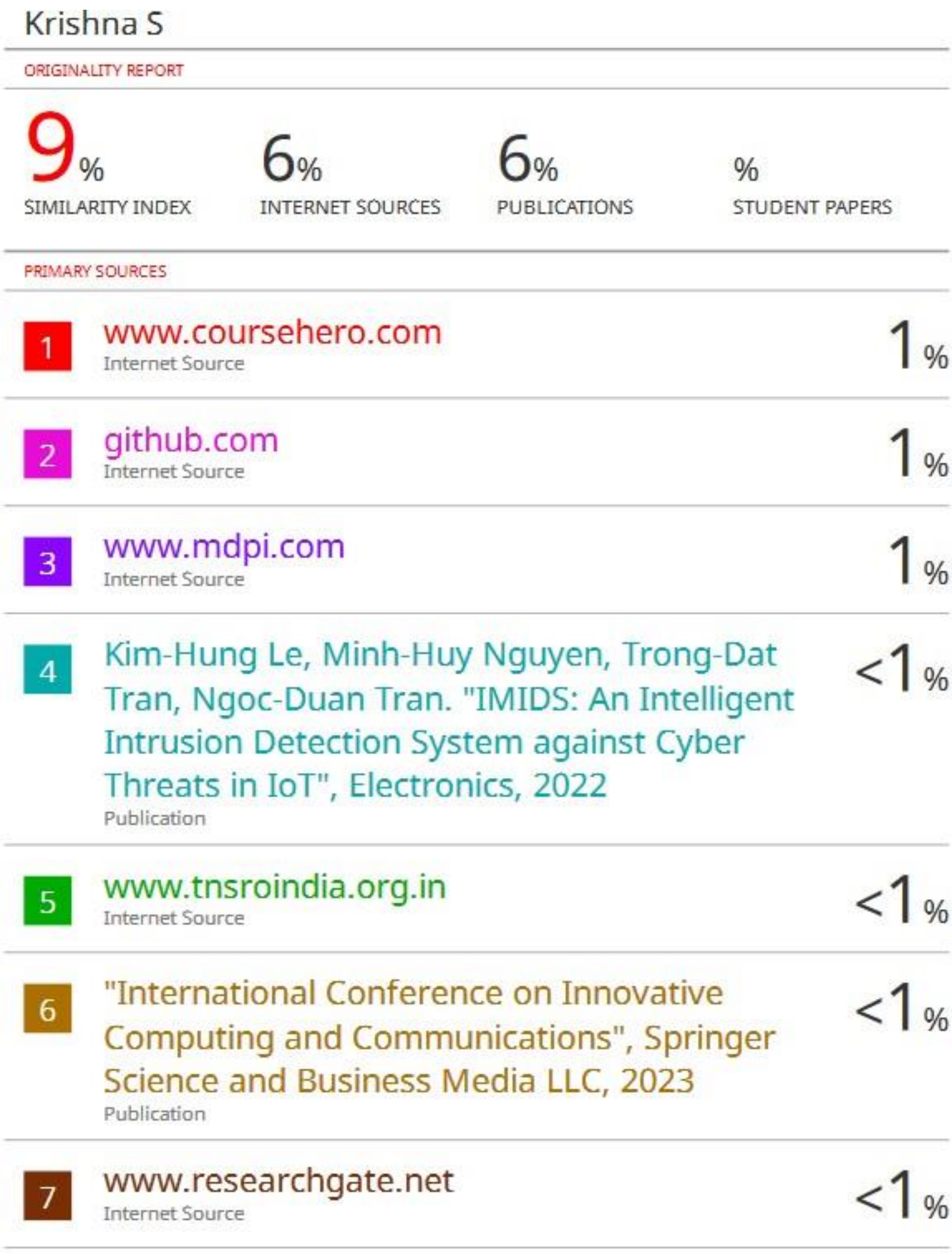
[31] Lee, Jonghoon & Kim, Jonghyun & Kim, Ikkyun & Han, Kijun. (2019). Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2953095.

[32] Siniosoglou, Ilias & Radoglou Grammatikis, Panagiotis & Efstathopoulos, George & Fouliras, Panagiotis & Sarigiannidis, Panagiotis. (2021). A Unified Deep Learning Anomaly Detection and Classification Approach for Smart Grid Environments. IEEE Transactions on Network and Service Management. PP. 10.1109/TNSM.2021.3078381.

[33] Li, Dan & Chen, Dacheng & Jin, Baihong & Shi, Lei & Goh, Jonathan & Ng, See-Kiong. (2019). MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks. 10.1007/978-3-030-30490-4_56.

[34] Gao, Ying & Liu, Yu & Jin, Yaqia & Chen, Juequan & Wu, Hongrui. (2018). A Novel Semi-Supervised Learning Approach for Network Intrusion Detection on Cloud-Based Robotic System. IEEE Access. 6. 1-1. 10.1109/ACCESS.2018.2868171.

[35] Piplai, Aritran & Chukkapalli, Sai & Joshi, Anupam. (2020). NAttack! Adversarial Attacks to bypass a GAN based classifier trained to detect Network intrusion.



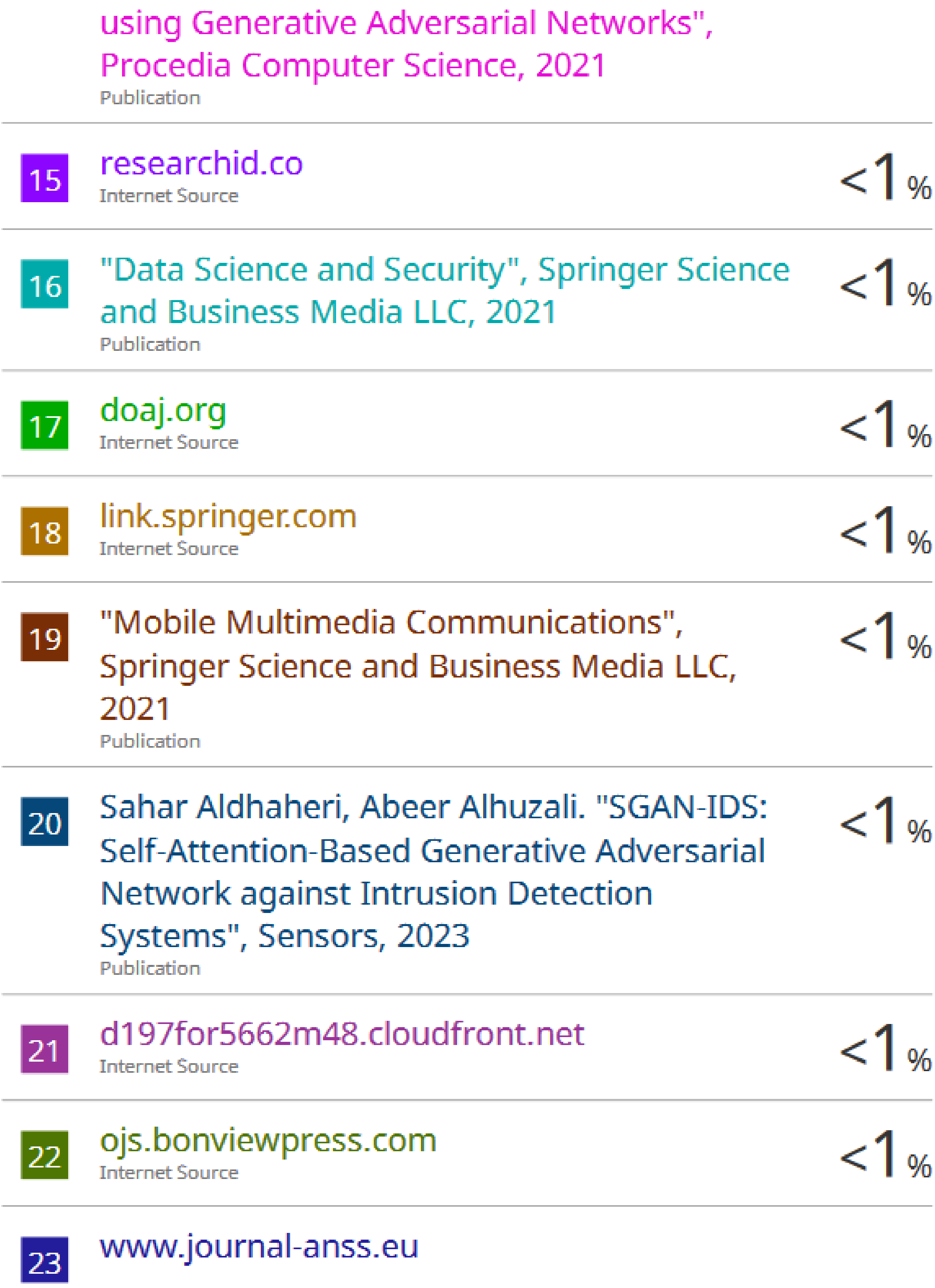
Krishna S

ORIGINALITY REPORT

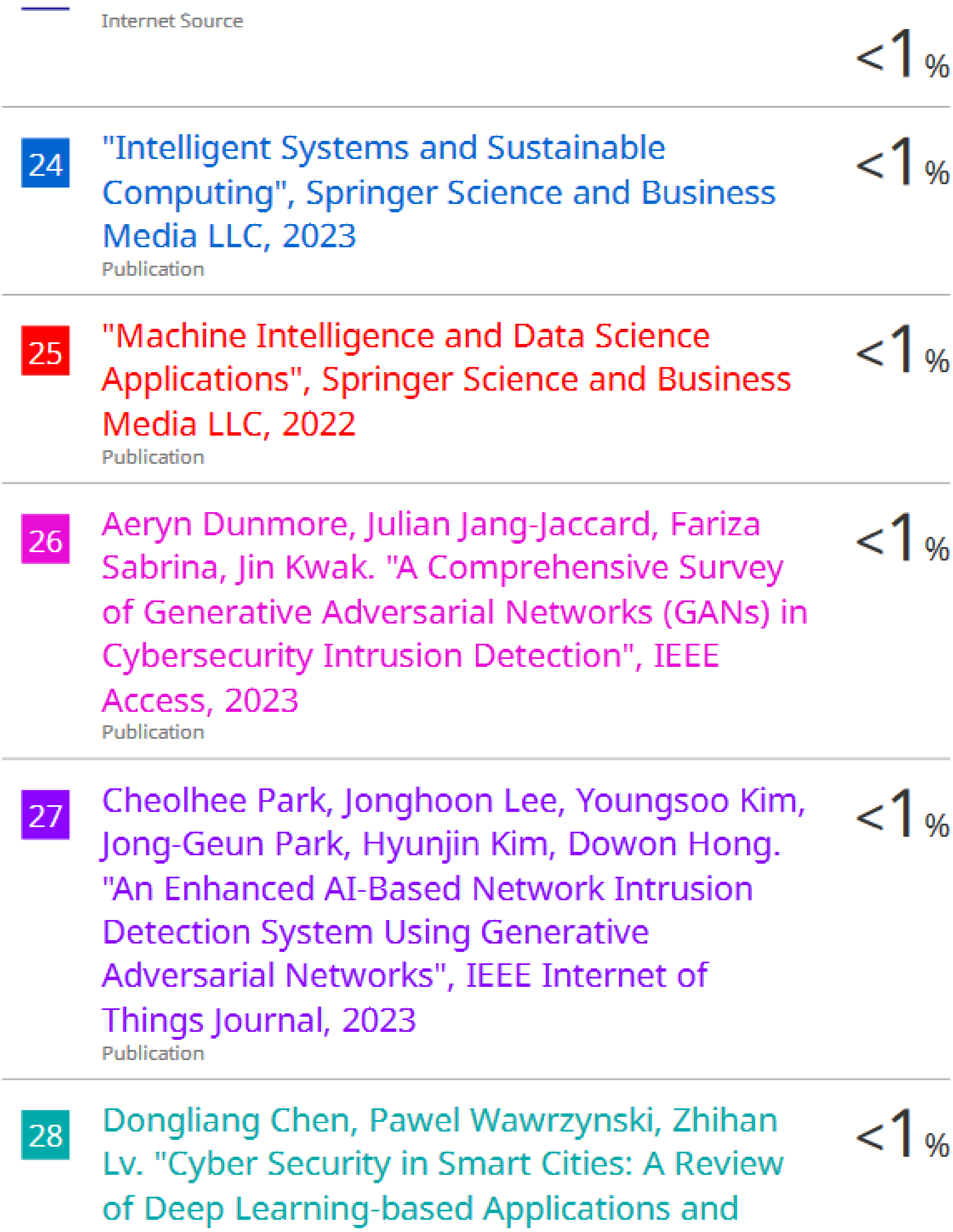
9%	6%	6%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

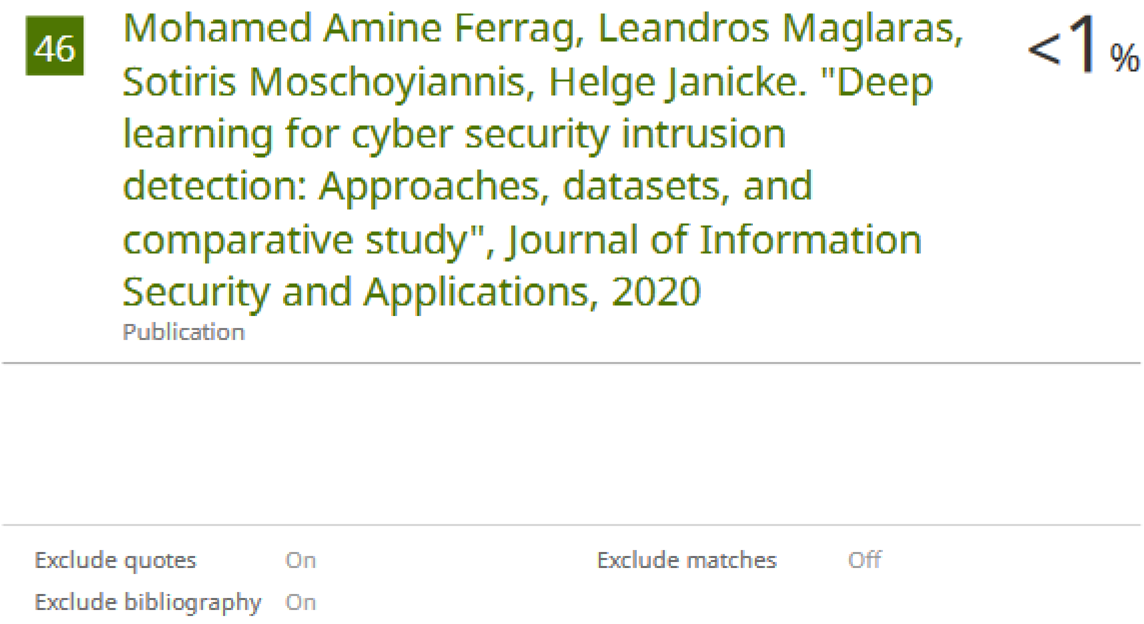
1	www.coursehero.com Internet Source	1%
2	github.com Internet Source	1%
3	www.mdpi.com Internet Source	1%
4	Kim-Hung Le, Minh-Huy Nguyen, Trong-Dat Tran, Ngoc-Duan Tran. "IMIDS: An Intelligent Intrusion Detection System against Cyber Threats in IoT", Electronics, 2022 Publication	<1%
5	www.tnsroindia.org.in Internet Source	<1%
6	"International Conference on Innovative Computing and Communications", Springer Science and Business Media LLC, 2023 Publication	<1%
7	www.researchgate.net Internet Source	<1%



	using Generative Adversarial Networks", Procedia Computer Science, 2021 Publication	
15	researchid.co Internet Source	<1 %
16	"Data Science and Security", Springer Science and Business Media LLC, 2021 Publication	<1 %
17	doaj.org Internet Source	<1 %
18	link.springer.com Internet Source	<1 %
19	"Mobile Multimedia Communications", Springer Science and Business Media LLC, 2021 Publication	<1 %
20	Sahar Aldhaheri, Abeer Alhuzali. "SGAN-IDS: Self-Attention-Based Generative Adversarial Network against Intrusion Detection Systems", Sensors, 2023 Publication	<1 %
21	d197for5662m48.cloudfront.net Internet Source	<1 %
22	ojs.bonviewpress.com Internet Source	<1 %
23	www.journal-anss.eu	



	Internet Source	<1 %
24	"Intelligent Systems and Sustainable Computing", Springer Science and Business Media LLC, 2023 Publication	<1 %
25	"Machine Intelligence and Data Science Applications", Springer Science and Business Media LLC, 2022 Publication	<1 %
26	Aeryn Dunmore, Julian Jang-Jaccard, Fariza Sabrina, Jin Kwak. "A Comprehensive Survey of Generative Adversarial Networks (GANs) in Cybersecurity Intrusion Detection", IEEE Access, 2023 Publication	<1 %
27	Cheolhee Park, Jonghoon Lee, Youngsoo Kim, Jong-Geun Park, Hyunjin Kim, Dowon Hong. "An Enhanced AI-Based Network Intrusion Detection System Using Generative Adversarial Networks", IEEE Internet of Things Journal, 2023 Publication	<1 %
28	Dongliang Chen, Pawel Wawrzynski, Zhihan Lv. "Cyber Security in Smart Cities: A Review of Deep Learning-based Applications and	<1 %



46 Mohamed Amine Ferrag, Leandros Maglaras, Sotiris Moschoyiannis, Helge Janicke. "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study", Journal of Information Security and Applications, 2020 <1%

Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		