International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. #, No.# (05-2024)

Using Artificial Intelligence to detect evasive techniques in Contemporary technologies

Mohammed Almuqrin[1], Shailendra Mishra[2]
College of Computer Sciences and Information Technology, Majmaah University, Majmaah, Saudi Arabia, 441104797@s.mu.edu.sa[1]
Department of Computer Engineering, College of Computer and Information Sciences, Majmaah University, Majmaah, Saudi Arabia, s.mishra@mu.edu.sa[2]

***Abstract*— This study focuses on the design of an artificial intelligence (AI) tool dedicated to monitoring and distinguishing secure and insecure communication flows within a company. The primary objective is to ensure the secure transfer of data by constructing a cyber-secure model using AI. The methodology involves training the model on a comprehensive database encompassing various communication protocols within the company, including both benign and malicious communications. The background emphasizes the significance of safeguarding company communications and data transfer, setting the stage for the study's purpose. In terms of methods, the project employs AI techniques to build a cyber-secure model capable of discerning the security status of communication channels. The model is trained on a diverse dataset covering all communication protocols utilized within the company, ensuring its adaptability to various scenarios. The focus on AI-driven security sets the project apart in addressing contemporary challenges in data protection. Results from the study highlight the successful development and training of the AI model, showcasing its ability to distinguish between secure and insecure communication channels. The model's effectiveness is demonstrated through its comprehensive understanding of different communication protocols, enabling it to accurately identify and classify secure and insecure data transfers. Conclusions drawn from the study emphasize the pivotal role of AI in enhancing cybersecurity within corporate networks. The successful implementation of the AI tool provides a proactive approach to identifying and securing communication flows, mitigating potential risks associated with insecure data transfer. The study underscores the potential of AI-driven solutions in fortifying cyber defenses and ensuring the integrity of communication within organizational frameworks. In summary, AI tools have emerged as a robust and effective means to bolster the security of company communications, contributing to the ongoing efforts to safeguard sensitive data in corporate environments.**

Keywords—Artificial Intelligence Security; Cybersecurity Monitoring; Secure Data Transfer; Communication Protocols Classification; Corporate Network Cyber Defense

I. Introduction

in the ever-evolving landscape of cybersecurity, the development of robust and intelligent systems is paramount to safeguarding digital environments from the incessantly mutating threat landscape. This article introduces a groundbreaking project focused on conceiving and deploying an artificial intelligence (AI) system tailored for cybersecurity, with a primary emphasis on harnessing the UNSW-NB15 dataset. This dataset, originating from the esteemed Cyber Range Lab of the Australian Centre for Cyber Security (ACCS), represents a unique amalgamation of meticulously crafted raw network packets generated by the IXIA PerfectStorm tool. Comprising real-world modern activities and synthetically generated contemporary attack behaviors, the UNSW-NB15 dataset provides a comprehensive foundation for training and testing AI models [1].

To unlock the full potential of the UNSW-NB15 dataset, the Tcpdump tool captures a substantial 100 GB of raw traffic, resulting in Pcap files encompassing nine distinct types of attacks. These attacks range from Fuzzers and Analysis to Backdoors, Denial of Service (DoS), Exploits, Generic threats, Reconnaissance, Shellcode, and Worms. The integration of advanced tools such as Argus and Bro-IDS, coupled with the development of twelve specialized algorithms, culminates in the extraction of 49 features, including a class label meticulously detailed in the UNSW-NB15_features.csv file [2].

With a staggering two million and 540,044 records, the dataset is distributed across four CSV files (UNSW-NB15_1.csv, UNSW-NB15_2.csv, UNSW-NB15_3.csv, and UNSW-NB15_4.csv). Enhancing its comprehensiveness, the ground truth table (UNSW-NB15_GT.csv) and event list file (UNSW-NB15_LIST_EVENTS.csv) further enrich the dataset. A thoughtful partitioning strategy creates a training set (UNSW_NB15_training-set.csv) with 175,341 records and a testing set (UNSW_NB15_testing-set.csv) containing 82,332 records, encompassing various attack types and normal behaviors.

In the dynamic landscape of cybersecurity, the imperative to develop robust and intelligent systems is more crucial than ever to protect digital environments from the incessantly mutating threat landscape. This article delves into a pioneering project centered around the conception and deployment of an artificial intelligence (AI) system tailored for cybersecurity, with a focal point on harnessing the UNSW-NB15 dataset. This dataset,

E-mail: 441104797@s.mu.edu.sa

originating from the esteemed Cyber Range Lab of the Australian Centre for Cyber Security (ACCS), represents a unique amalgamation of meticulously crafted raw network packets generated by the IXIA PerfectStorm tool. Comprising real-world modern activities and synthetically generated contemporary attack behaviors, the UNSW-NB15 dataset provides a comprehensive foundation for training and testing AI models [1],[2].

Amidst the evolving challenges of the digital age, the importance of artificial intelligence in cybersecurity cannot be overstated. The implementation of intelligent systems not only fortifies defenses against a diverse array of cyber threats but also offers substantial economic and temporal advantages. Unlike traditional security measures that may necessitate extensive human resources, an AI-driven approach streamlines the process, potentially saving significant costs and time associated with manual intervention [3].

The UNSW-NB15 dataset, with its blend of genuine modern network activities and synthetic attack behaviors, serves as an ideal playground for the development and evaluation of advanced AI models. The integration of the tcpdump tool, capturing a substantial 100 GB of raw traffic, enables the creation of Pcap files encompassing nine distinct attack types. These attacks range from Fuzzers and Analysis to Backdoors, Denial of Service (DoS), Exploits, Generic threats, Reconnaissance, Shellcode, and Worms. The subsequent utilization of advanced tools such as Argus and Bro-IDS, along with the implementation of twelve specialized algorithms, results in the extraction of 49 features, providing a rich dataset with a detailed class label in UNSW-NB15_features.csv.

The sheer volume of data, totaling two million and 540,044 records across four CSV files, namely UNSW-NB15_1.csv, UNSW-NB15_2.csv, UNSW-NB15_3.csv, and UNSW-NB15_4.csv, enhances the comprehensiveness of the dataset. The ground truth table (UNSW-NB15_GT.csv) and event list file (UNSW-NB15_LIST_EVENTS.csv) further enrich the dataset, contributing to its credibility and applicability.

One of the prime advantages of implementing AI systems in cybersecurity lies in their ability to provide continuous, real-time monitoring of a company's network. This not only ensures a proactive defense against potential threats but also allows network managers to receive instantaneous indicators of communication status. Whether online or offline, the applied model empowers network managers by offering insights into the communications within the organization, along with an assessment of potential risks. In cases where the risk threshold is exceeded, the system can autonomously initiate preventive measures, potentially thwarting an ongoing cyber attack [4].

As the study unfolds, it aspires to contribute novel insights and methodologies, fortifying the capabilities of AI systems in countering cyber threats. The introduction, as presented, provides a comprehensive overview, making the intricate subject matter accessible to scientists beyond the specific domain. It emphasizes the project's main aims and principal conclusions while highlighting the transformative potential of AI in reshaping the landscape of cybersecurity defenses [3].

This study aims to harness the vast information encapsulated in the UNSW-NB15 dataset to train an AI model, elevating cybersecurity defenses to new heights. The introduction contextualizes the project within the broader scope of cybersecurity research, emphasizing its importance in addressing the constantly shifting threat landscape. It defines the purpose of the work, underscoring its significance in advancing the field of AI-driven cybersecurity [4].

In the cybersecurity landscape, sophisticated and elusive threats pose significant challenges, this research introduces an approach utilizing advanced artificial intelligence (AI) to redefine digital forensics. With a firm commitment to ethical principles, including transparency, privacy protection, and bias minimization, we aim to build a foundation of trust and responsibility. This ethical framework guides the development of AI systems, ensuring their alignment with core values while enhancing the detection and response capabilities against cyber threats.

A. *The objectives of this study are;*

i. Creating an AI-driven framework tailored to identify and adapt to emerging cyber threats, enhancing the precision and adaptability of cybersecurity measures.

ii. Designing AI algorithms with advanced learning capabilities to uncover previously undetectable threats, pushing the boundaries of traditional security systems.

iii. Implementing systems that provide clear, actionable insights from security alerts, streamlining cybersecurity professionals' response processes.

iv. Enhancing decision-making in cybersecurity operations with AI-derived insights, supporting a strategic approach to threat mitigation.

v. Ensuring the ethical deployment of AI, focusing on reducing biases, safeguarding privacy, and securing data to cultivate trust among stakeholders.

vi. Validating the AI model's effectiveness through rigorous testing against real-world scenarios, confirming its readiness for deployment in cybersecurity contexts.

Through these initiatives, the research strives to significantly advance the field of cybersecurity digital forensics, offering a more effective, reliable, and ethically responsible approach to navigating the complex landscape of cyber threats.

E-mail: 441104797@s.mu.edu.sa

As the project unfolds, it aspires to contribute novel insights and methodologies, fortifying the capabilities of AI systems in countering cyber threats. In conclusion, the introduction provides a comprehensive overview, making the intricate subject matter accessible to scientists beyond the specific domain, while emphasizing the project's main aims and principal conclusions [5].

Despite the comprehensive nature of the UNSW-NB15 dataset and its applications in cybersecurity research, there remain gaps in the current state of network intrusion detection systems (NIDS), particularly in the detection and classification of sophisticated and novel attack vectors. Existing systems may lack adaptability and the ability to scale with the continually evolving threat landscape, often resulting in high false-positive rates. My study addresses these gaps by employing an AI-driven approach that not only encompasses traditional attack signatures but also employs machine learning techniques to learn from network traffic patterns. This allows for a more dynamic and proactive response to emerging threats.

II. Related Work

The study "AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling, and Research Directions" explores the transformative impact of Artificial Intelligence (AI) in the context of the Fourth Industrial Revolution, often referred to as Industry 4.0. Positioned as a pivotal force, the paper underscores the role of AI in safeguarding Internet-connected systems from a myriad of cyber threats, damages, and unauthorized access [6]. A significant aspect of the study is its emphasis on the practical application of AI techniques, specifically machine learning and deep learning, in enhancing cybersecurity measures. The paper advocates for an AI-driven Cybersecurity approach, emphasizing the automation and intelligence inherent in security intelligence modeling. This paradigm shift from conventional security systems is considered a substantial advancement in the field.

Beyond a retrospective analysis, the study outlines potential research directions, serving as a roadmap for future endeavors in the domain. The overarching goal is to provide comprehensive guidelines for both cybersecurity researchers and industry professionals. By advocating for an intelligent computing or AI-based technical viewpoint, the study aims to facilitate advancements in the application of AI for cybersecurity.

The study [7] delves into the transformative influence of AI on cybersecurity, recognizing it as a catalyst for value creation across diverse sectors. By focusing on AI's broad applications, particularly in various business sectors, the paper highlights its relevance in addressing the evolving challenges in cybersecurity due to increased reliance on information technology [7]. Employing a literature review approach, the study delves into existing research, providing a comprehensive understanding of the multifaceted impacts of AI on cybersecurity. The research identifies AI as a major influencer on a large scale, driven by the imperative need for enhanced security measures in recent technologies.

The study is encapsulated by keywords such as Cybersecurity, AI, Cyber threats, Vulnerability, Data privacy, and AI value creation. These terms embody the essence of the research, highlighting the symbiotic relationship between AI and cybersecurity. The research paper positions AI as a transformative force, not only enhancing cybersecurity but also generating value across various domains.

Centered around the role of AI in the cybersecurity market, this study provides a comprehensive scheme to assist organizations in countering cyber threats. It emphasizes the purpose of AI in enhancing organizations' capabilities to monitor and safeguard data, responding to the rising global awareness of cyber threats and the need for robust information security [8]

The research explores driving factors such as increased awareness, technological advancements, and intelligence system upgrades, acknowledging the surge in data from diverse sources. Insights from previous research related to AI and cybersecurity are incorporated, positioning AI as a pivotal element in fortifying cybersecurity measures amid escalating threats.

The study [8] recognizes various motivations behind cyber-attacks, including political rivalry, competitive motives, reputation damage, international data theft, and interests of radical religious groups. The acknowledgment of these motivations underscores the need for advanced, intelligent, and adaptive AI-driven solutions in the modern cybersecurity paradigm.

Delving into the expansive realm of cybersecurity, this study emphasizes the diverse applications of AI and Machine Learning (ML) across organizations and governments. The paper highlights the role of AI and ML in fortifying security measures, with a specific focus on digital transactions [9].

The study[9] underscores the synergy between cybersecurity and AI, showcasing the effectiveness of concurrent integration. ML approaches stand out by addressing the limitations of earlier rule-based security structures, providing a more robust and dynamic security framework. The study concludes by emphasizing the evolving landscape of cybersecurity, where the strategic integration of AI and ML technologies proves to be a beacon of progress, offering current trends and prospects in the field.

In the paper [10], "Artificial Intelligence in Cybersecurity: The Use of AI Along the Cyber Kill Chain," authored by Iwona Chomiak-Orsa, Artur Rot, and Bartosz Blaicke and published on August 9, 2019, the authors address the contemporary challenge in defending against cyberattacks, highlighting that the speed and quantity of threats often surpass human-centered cyber defense capabilities. They advocate for an Artificial Intelligence

(AI)-driven approach to enhance the effectiveness of security controls while acknowledging the dual nature of AI, which can be exploited by adversaries to create more sophisticated attack mechanisms.

With dual objectives, the paper aims to identify where along the cyber kill chain AI capabilities have been applied and determine the phase with the greatest near-term potential, considering recent developments and publications. The authors focus on three key AI capabilities: knowledge acquisition, human-like perception, and decision-making, using them as the basis for evaluating AI's role in different phases of the cyber kill chain.

The study concludes by underlining the critical need for AI integration in cybersecurity to counter the evolving threat landscape. It provides insights into AI applications along the cyber kill chain, identifying areas with the greatest potential for future deployment. However, the authors emphasize the importance of cautious consideration, recognizing the risks associated with AI as a potential tool for adversaries in crafting advanced cyber threats.

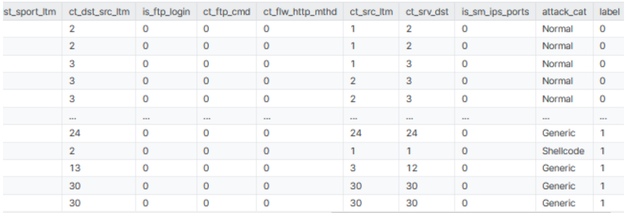
Despite the comprehensive nature of the UNSW-NB15 dataset and its applications in cybersecurity research, there remain gaps in the current state of network intrusion detection systems (NIDS), particularly in the detection and classification of sophisticated and novel attack vectors. Existing systems may lack adaptability and the ability to scale with the continually evolving threat landscape, often resulting in high false-positive rates. My study addresses these gaps by employing an AI-driven approach that not only encompasses traditional attack signatures but also employs machine learning techniques to learn from network traffic patterns. This allows for a more dynamic and proactive response to emerging threats.

III. Proposed Work

A. Data Set

In our study, we have embarked on the development of an artificial intelligence (AI) tool for monitoring network traffic to discern between secure and malicious communications. Our objective revolves around the meticulous design of an AI-driven system capable of identifying potential security threats within network flows. This undertaking aligns with the pressing need for advanced cybersecurity solutions, given the escalating frequency and sophistication of cyberattacks. The foundation of our work rests upon the utilization of a neural network, specifically a neural model trained on the UNSW-NB15 dataset, a comprehensive resource curated by Dr. Nour Moustafa and Jill Slay.

TABLE I. DataSet Content Description



st_sport_ltm	ct_dst_src_ltm	is_ftp_login	ct_ftp_cmd	ct_flw_http_mthd	ct_src_ltm	ct_srv_dst	is_sm_ips_ports	attack_cat	label
	2	0	0	0	1	2	0	Normal	0
	2	0	0	0	1	2	0	Normal	0
	3	0	0	0	1	3	0	Normal	0
	3	0	0	0	2	3	0	Normal	0
	3	0	0	0	2	3	0	Normal	0
	...	...	...	...	...	...	...	...	...
	24	0	0	0	24	24	0	Generic	1
	2	0	0	0	1	1	0	Shellcode	1
	13	0	0	0	3	12	0	Generic	1
	30	0	0	0	30	30	0	Generic	1
	30	0	0	0	30	30	0	Generic	1

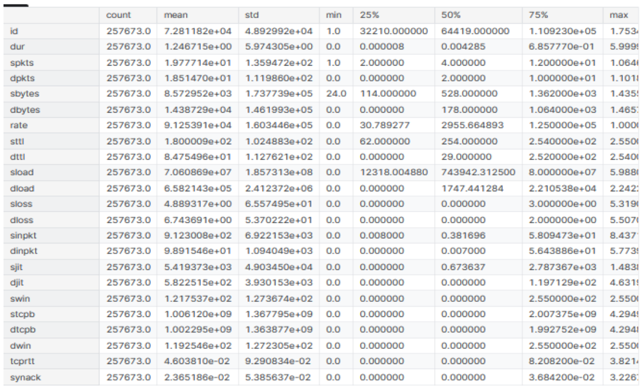
The UNSW-NB15 dataset is a comprehensive collection of raw network packets, meticulously curated in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS). Generated by the IXIA PerfectStorm tool, this dataset aims to provide a hybrid of genuine contemporary network activities and synthetic attack behaviors. Utilizing the Tcpdump tool, 100 GB of raw traffic data, in the form of Pcap files, was captured. This dataset encompasses nine distinct types of attacks, including Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. The Argus and Bro-IDS tools were employed, resulting in the creation of twelve algorithms that generated a total of 49 features, each labeled with a specific class. These features are detailed in the UNSW-NB15_features.csv file.

The UNSW-NB15 dataset is a comprehensive collection of raw network packets, meticulously curated in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS). Generated by the IXIA PerfectStorm tool, this dataset aims to provide a hybrid of genuine contemporary network activities and synthetic attack behaviors. Utilizing the Tcpdump tool, 100 GB of raw traffic data, in the form of Pcap files, was captured. This dataset encompasses nine distinct types of attacks, including Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. The Argus and Bro-IDS tools were employed, resulting in the creation of twelve algorithms that generated a total of 49 features, each labeled with a specific class. These features are detailed in the UNSW-NB15_features.csv file.

B. Data Set Processing

To uphold the principles of transparent and replicable research, we adhere to a rigorous description of our methods, as stipulated by the guidelines. This commitment ensures that our methodologies are sufficiently detailed, allowing fellow researchers to replicate and build upon our results. Furthermore, by the ethos of open science, we pledge to make all materials, data, computer code, and protocols associated with our publication readily available to readers. This includes the provision of comprehensive flow diagrams, flow charts, and pseudocode, essential for understanding the intricacies of our AI-driven network monitoring tool. In crafting our approach, we draw inspiration from the notable work of Dr. Nour Moustafa and Jill Slay, particularly their UNSW-NB15 dataset, which has been instrumental in advancing intrusion detection systems. The UNSW-NB15 dataset, with its diverse set of samples encompassing various types of malware and benign traffic, serves as a robust foundation for our research. We recognize the significance of distinguishing between well-established methods and new protocols in our description. As such, our article provides detailed insights into the novel techniques employed, while also giving appropriate citations for established methodologies. The methods employed in our project include a comprehensive series of steps aimed at achieving effective network traffic monitoring. We commence with a thorough analysis and exploration of the data, emphasizing the significance of understanding variables and formulating a modeling strategy [11,12,13]. Subsequently, data preparation involves restructuring the dataset to suit our classification needs, including binary and multi-class classifications. Our data preprocessing encompasses imputation and encoding, ensuring that the dataset is appropriately formatted for analysis.

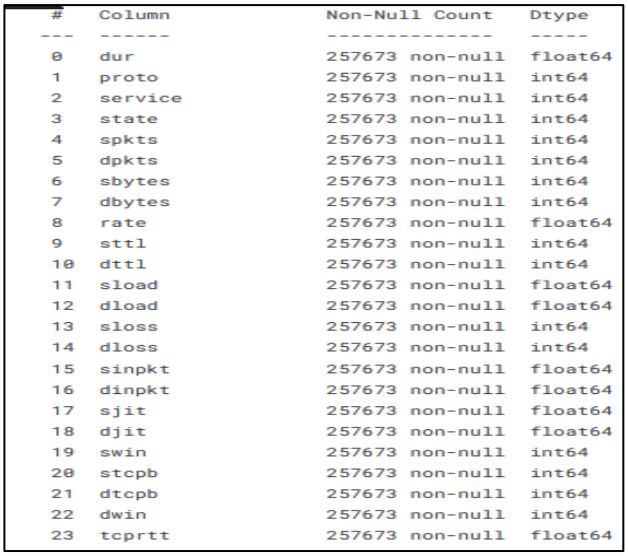
TABLE II. DATA PREPARATION FOR CLASSIFICATION



	count	mean	std	min	25%	50%	75%	max
id	257673.0	7.281182e+04	4.892992e+04	1.0	32210.000000	64419.000000	1.109230e+05	1.753
dur	257673.0	1.246715e+00	5.974305e+00	0.0	0.000008	0.004285	6.857770e-01	5.999
spkts	257673.0	1.977714e+01	1.359472e+02	1.0	2.000000	4.000000	1.200000e+01	1.064
dpkts	257673.0	1.851470e+01	1.119860e+02	0.0	0.000000	2.000000	1.000000e+01	1.101
sbytes	257673.0	8.572952e+03	1.737739e+05	24.0	114.000000	528.000000	1.362000e+03	1.435
dbytes	257673.0	1.438729e+04	1.461993e+05	0.0	0.000000	178.000000	1.064000e+03	1.465
rate	257673.0	9.125391e+04	1.603446e+05	0.0	30.789277	2955.664893	1.250000e+05	1.000
sttl	257673.0	1.800009e+02	1.024883e+02	0.0	62.000000	254.000000	2.540000e+02	2.550
dttl	257673.0	8.475496e+01	1.127621e+02	0.0	0.000000	29.000000	2.520000e+02	2.540
sload	257673.0	7.060869e+07	1.857313e+08	0.0	12318.004880	743942.312500	8.000000e+07	5.988
dload	257673.0	6.582143e+05	2.412372e+06	0.0	0.000000	1747.441284	2.210538e+04	2.242
sloss	257673.0	4.889317e+00	6.557495e+01	0.0	0.000000	0.000000	3.000000e+00	5.319
dloss	257673.0	6.743691e+00	5.370222e+01	0.0	0.000000	0.000000	2.000000e+00	5.507
sinpkt	257673.0	9.123008e+02	6.922153e+03	0.0	0.008000	0.381696	5.809473e+01	8.437
dinpkt	257673.0	9.891546e+01	1.094049e+03	0.0	0.000000	0.007000	5.643886e+01	5.773
sjit	257673.0	5.419373e+03	4.903450e+04	0.0	0.000000	0.673637	2.787367e+03	1.483
djit	257673.0	5.822515e+02	3.930153e+03	0.0	0.000000	0.000000	1.197129e+02	4.631
swin	257673.0	1.217537e+02	1.273674e+02	0.0	0.000000	0.000000	2.550000e+02	2.550
stcpb	257673.0	1.006120e+09	1.367795e+09	0.0	0.000000	0.000000	2.007375e+09	4.294
dtcpb	257673.0	1.002295e+09	1.363877e+09	0.0	0.000000	0.000000	1.992752e+09	4.294
dwin	257673.0	1.192546e+02	1.272305e+02	0.0	0.000000	0.000000	2.550000e+02	2.550
tcprtt	257673.0	4.603810e-02	9.290834e-02	0.0	0.000000	0.000000	8.208200e-02	3.821
synack	257673.0	2.365186e-02	5.385637e-02	0.0	0.000000	0.000000	3.684200e-02	3.226

Table II shows the database after it has been processed, and cleaned, and zero and empty values have been removed, making it ready for training. The subsequent steps involve normalization, feature extraction, and model architecture design. These steps are critical in enhancing the quality of data and facilitating effective learning by the AI algorithms. For feature extraction, we employ embedded methods utilizing the random forest algorithm to determine the relevance of attributes. The selected attributes are then used to train our AI models. The imputation process addresses missing or infinite values in the dataset, ensuring its integrity. In encoding, we label our data to represent different classes, and in normalization, we scale variables to a uniform magnitude for optimal model training. Our commitment to transparency extends to the provision of detailed data repartition information, showcasing the distribution of the dataset across training and testing sets for both binary and multi-class classifications.

TABLE III. DATA SET FOR CLEANING



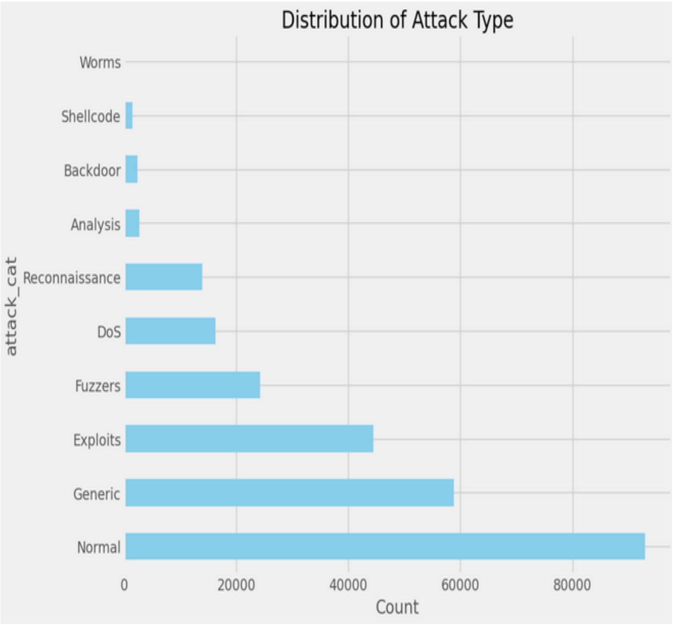
```
 #   Column   Non-Null Count   Dtype
---  ------   --------------   -----
 0   dur      257673 non-null  float64
 1   proto    257673 non-null  int64
 2   service  257673 non-null  int64
 3   state    257673 non-null  int64
 4   spkts    257673 non-null  int64
 5   dpkts    257673 non-null  int64
 6   sbytes   257673 non-null  int64
 7   dbytes   257673 non-null  int64
 8   rate     257673 non-null  float64
 9   sttl     257673 non-null  int64
 10  dttl     257673 non-null  int64
 11  sload    257673 non-null  float64
 12  dload    257673 non-null  float64
 13  sloss    257673 non-null  int64
 14  dloss    257673 non-null  int64
 15  sinpkt   257673 non-null  float64
 16  dinpkt   257673 non-null  float64
 17  sjit     257673 non-null  float64
 18  djit     257673 non-null  float64
 19  swin     257673 non-null  int64
 20  stcpb    257673 non-null  int64
 21  dtcpb    257673 non-null  int64
 22  dwin     257673 non-null  int64
 23  tcprtt   257673 non-null  float64
```

Table III Visualizing our target variables further aids in understanding the dataset, highlighting its unbalanced nature. To mitigate bias in our models, we employ oversampling through the SMOTE (synthetic minority over-sampling technique) Python library, addressing the data imbalance issue.

C. Data Set Classification

The architecture of our models encompasses both deep learning and machine learning approaches. For deep learning, we implement Convolutional Neural Network (CNN) and Deep Neural Network (DNN) models. The CNN model, typically used for attack detection, is adapted for malware classification, with an emphasis on parameter adjustments for optimal results [14]. The DNN model, known for its efficiency in obtaining data representation, is structured with input, intermediate, and output layers. Additionally, machine learning algorithms, specifically the Random Forest (RF) classifier and Decision Tree (DT) classifier, are incorporated into our model architecture. These classifiers contribute to the predictive capabilities of our system, aiding in the classification of malware and benign traffic as shown in Fig I.

FIG I. DISTRIBUTION OF ATTACK TYPES

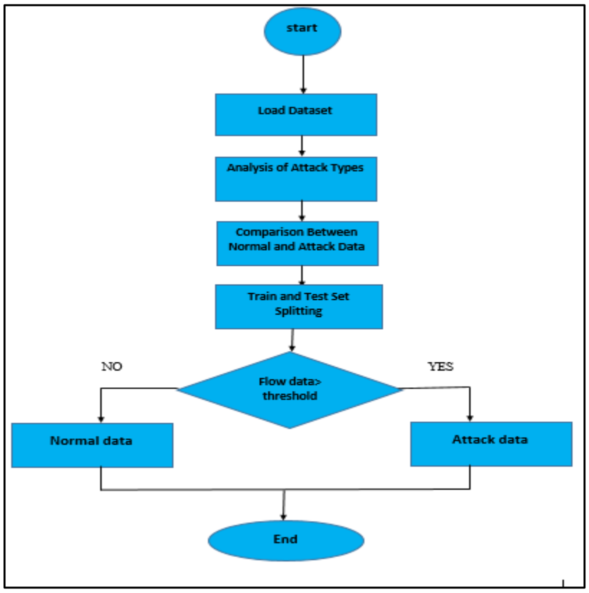


Throughout our project, we ensure that each step is meticulously explained, enabling researchers to replicate and build upon our findings. The use of diagrams, such as the architecture illustrations for attacks, adds a visual dimension to our methodology, enhancing comprehension.

IV. Experimental Setup

The following flow chart (Fig II) shows the mechanism of operation of the neural network on which the artificial intelligence model that we designed in this research depends, indicating the work steps.

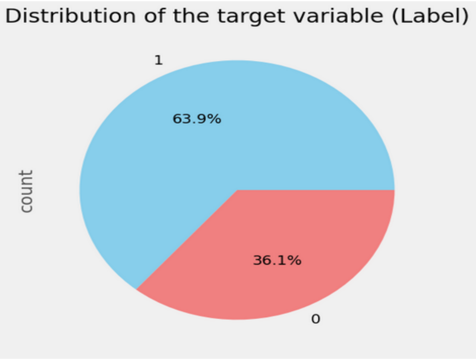
FIG II. MECHANISM OF OPERATION OF THE NEURAL NETWORK



V. Results & Discussion

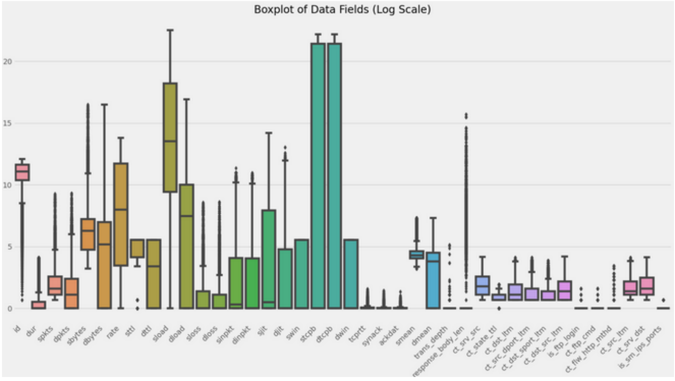
Previous studies have laid a solid foundation for understanding the use of AI for cybersecurity. However, they often do not provide a solution that evolves with the threat landscape. Through the research, it’s observed that by applying a machine learning model trained on a broad dataset encompassing a wide range of attack types, it is possible to reduce false positives and adapt to new threats more effectively. This insight is crucial as it demonstrates the capability of AI not just to detect known threats but also to adapt and respond to new patterns of attacks, which has been a limitation in many existing systems. In this section, we visualized and understood the distribution of the target variable (label) in the dataset. We utilized the value_counts() code to calculate the occurrences of each label. Subsequently, a pie chart was created using the plot() function with the 'pie' kind to represent the distribution visually. The chart was customized with colors for better clarity, and the percentages of each label were displayed using the autopct parameter. The resulting visualization provides an insightful representation of the distribution of labels in the dataset, aiding in the initial understanding of the data's composition as shown in Fig III.

FIG III. DISTRIBUTION OF NORMAL AND ABNORMAL TRAFFIC



In this analysis, we employed a systematic approach to examine the numeric features within the dataset without explicitly using the term "code." Initially, we identified the numeric columns by selecting those with data types 'int64' and 'float64.' Subsequently, we created a boxplot visualization to gain insights into the distribution and potential outliers of these numeric variables. The plt. Fig(figsize=(20, 10)) command sets the size of the resulting plot to enhance visibility. We then utilized the Seaborn library's boxplot function to generate the boxplot, where the logarithmic transformation (np.log1p) was applied to the numeric columns. This transformation helps mitigate the impact of extreme values, making the distribution more interpretable. To enhance the readability of the plot, we adjusted the x-axis labels by rotating them 45 degrees to the right. The final touch involved adding a title to the plot, summarizing its purpose as "Boxplot of Data Fields (Log Scale)." This visual representation aids in identifying patterns, central tendencies, and potential outliers in the dataset's numeric features.

FIG IV. ANALYSIS OF DATA DISTRIBUTE

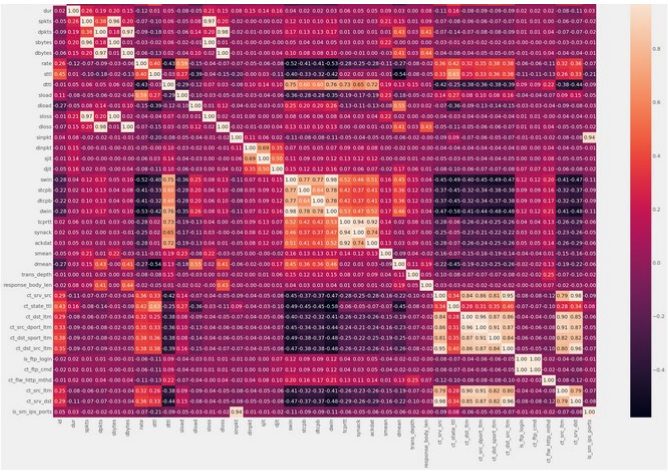


Initially, a correlation matrix is computed using the numeric columns from the EDA (Exploratory Data Analysis) data frame. This matrix represents the pairwise correlations between different numeric features in the dataset. Following the computation of the correlation matrix, a heatmap visualization is generated using the Seaborn library. The heatmap visually represents the correlation values, with annotations showing the correlation coefficients. The fmt=".2f" argument ensures that the correlation values are displayed with two decimal places for clarity. The plot is customized further with a figsize of (30, 20) to enhance visibility. The 'annot=True' parameter includes the numerical values in each cell of the heatmap, providing more detailed information. The title 'Correlation' is added to the plot to succinctly convey the purpose of visualizing the correlation matrix. This correlation analysis aids in understanding the relationships between different numeric features, helping identify patterns and dependencies in the data. Correlations closer to 1 or -1 indicate a stronger positive or negative relationship, respectively, while values closer to 0 signify a weaker or no correlation. In the context of the specified attack classifications:

Unusual Network Activity: High positive or negative correlations may indicate associations between unusual network activity and certain numeric features.Large Transactions: Correlations may reveal connections between large transaction characteristics and other numeric attributes.Error Feed: Patterns of errors or unrealistic values might be reflected in the correlation matrix [15,16].

Cyber Attacks: Correlation analysis can unveil relationships between cyber attacks and numeric features exhibiting unusual behavior.Measurement Bias: Correlations may point out connections between measurement bias and certain numeric variables.System Problems: Correlations might reveal associations between system problems and anomalous patterns in numeric features. This analytical approach contributes to a comprehensive understanding of potential relationships between various aspects of network data and the specified attack classifications [17] as shown in Fig V.

FIG V. CORRELATION OF DATA FLOW TYPES



A. Discussion

Fig V shows the process of splitting the dataset into training and testing sets, a crucial step in machine learning model development. The `train_test_split` function from the `sklearn.model_selection` module is utilized for this purpose as following steps:

i. Define Features and Target Variable (X and y)

-X represents the features, and it is obtained by dropping the 'label' column (target variable) from the shuffled DataFrame (shuffled_df).

- y is the target variable and consists of the values from the 'label' column.

ii. Split the Data

-train_test_split is employed to split the data into training and testing sets.

iii. The parameters are

X: Features

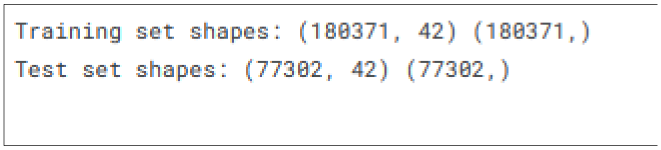
y: Target variable

-test_size=0.3: The proportion of the dataset to include in the test split. Here, 30% of the data is allocated for testing.

-random_state=42: Ensures reproducibility by setting a fixed random seed.

-stratify=y: Maintains the distribution of the target variable in both the training and testing sets, preserving the proportion of different classes.

FIG VI. TRAIN &TEST SET



```
Training set shapes: (180371, 42) (180371,)
Test set shapes: (77302, 42) (77302,)
```

This article involves the implementation of machine learning models, specifically a Random Forest classifier, a Multi-Layer Perceptron (MLP), and a Long Short-Term Memory (LSTM) network. The models are trained and evaluated for a dataset related to network security. Additionally, a Bayesian Gaussian Mixture Model (Beta Mixture Model) is applied for anomaly detection using the Beta distribution. The code performs cross-validation, evaluates ROC AUC (Receiver Operating Characteristic Area Under the Curve) for each fold, and determines an optimal threshold for anomaly detection.Random Forest, MLP, and LSTM Models:

i. -The Random Forest classifier is trained and evaluated using features and labels. Model performance metrics such as accuracy, F1 score, precision, and recall are computed.

ii. -The MLP model is constructed with multiple dense layers, and its training involves minimizing the sparse categorical cross-entropy loss. The model is then evaluated, considering metrics such as accuracy, loss, F1 score, precision, and recall.

iii. -The LSTM network is designed for sequence-based data, specifically with a binary classification task. It is trained using binary cross-entropy loss and evaluated for accuracy, loss, F1 score, precision, and recall.

iv. Beta Mixture Model for Anomaly Detection:

v. -A Bayesian Gaussian Mixture Model is employed for anomaly detection, specifically using the Beta distribution. The model is trained on a subset of the data labeled as normal. Anomaly scores are computed for the entire dataset.

vi. -Stratified K-Fold cross-validation is applied, and the ROC AUC is calculated for each fold. The threshold maximizing Youden's Index (sensitivity + specificity - 1) is determined for each fold.

vii. -The ROC curve is plotted, and the average ROC AUC and optimal threshold across all folds are computed.

The results of each model and the Beta Mixture Model for anomaly detection should be interpreted in the context of network security. Authors should consider the implications of the models' performance metrics and how they align with the objectives of intrusion detection. Comparisons with previous studies and existing methodologies can be discussed. Limitations and challenges encountered during model development and evaluation should also be addressed.

As shown in Fig VI, The threshold selection process involves evaluating the performance of the anomaly detection model at different threshold levels. In the context of your code, three thresholds are considered: the default threshold (0.5), the average optimal threshold obtained through cross-validation, and the best threshold determined based on the highest ROC AUC.

iv. Default Threshold (0.5)

Predictions are made using a default threshold of 0.5 [18].

The resulting binary predictions are then evaluated for accuracy, precision, recall, and F1-score using ground truth labels.

v. Average Optimal Threshold

The average optimal threshold, obtained through cross-validation, is applied to generate binary predictions [19,20,21].

Performance metrics (accuracy, precision, recall, F1-score) are computed based on these predictions and the ground truth labels.

vi. Best Threshold (Based on ROC AUC)

The threshold that maximizes the ROC AUC is determined during cross-validation[22].

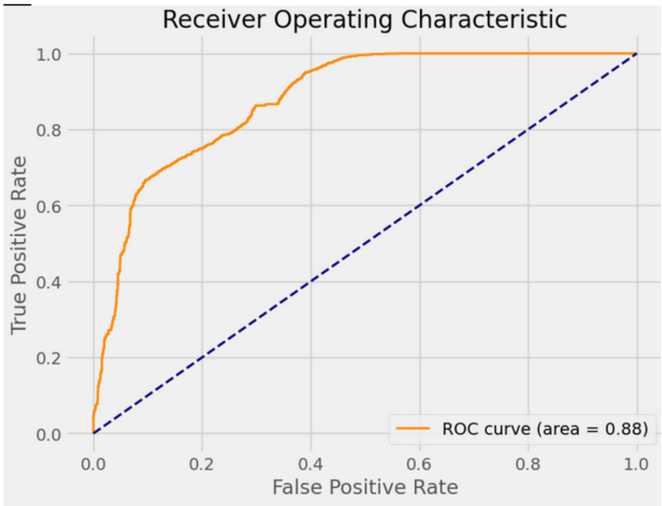
Binary predictions are made using this best threshold, and performance metrics are calculated accordingly.

vii. Performance Metrics Evaluation

For each threshold (default, average optimal, and best), binary predictions are compared with true labels to compute metrics.

Metrics include accuracy (overall correctness), precision (positive predictive value), recall (sensitivity or true positive rate), and F1-score (harmonic mean of precision and recall).

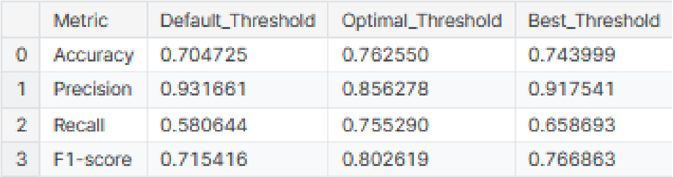
FIG VII. THRESHOLD FOR DEFINING ATTACKS



B. Evaluation Results

The results are organized in a data frame, summarizing the metrics for each threshold type as follows in Table 4.

TABLE IV. THRESHOLD VALUES



	Metric	Default_Threshold	Optimal_Threshold	Best_Threshold
0	Accuracy	0.704725	0.762550	0.743999
1	Precision	0.931661	0.856278	0.917541
2	Recall	0.580644	0.755290	0.658693
3	F1-score	0.715416	0.802619	0.766863

i. Threshold Significance

- The default threshold serves as a baseline, while the average optimal and best thresholds aim to enhance anomaly detection accuracy [22].
- Evaluation across multiple thresholds provides insights into how different threshold values impact model performance.

ii. -Optimal and best thresholds, especially derived from ROC analysis, offer a more nuanced understanding of model behavior in balancing true positive and false positive rates.

In summary, the threshold selection process involves assessing model performance at various threshold levels, enabling the identification of an optimal threshold for anomaly detection based on the chosen evaluation metrics. This approach ensures a thorough examination of the model's ability to distinguish between normal and anomalous instances in the dataset.

The limitations of the research may include potential biases in the dataset used, the reliance on specific algorithms that might not generalize well to diverse scenarios, and the dynamic nature of cyber threats, which could impact the model's adaptability over time. Additionally, the research findings may be influenced by the chosen features and may not fully capture emerging or novel attack patterns. Addressing these limitations could enhance the robustness and applicability of the research outcomes.

VI. CONCLUSIONS & FUTURE WORK

In conclusion, this research has provided valuable insights into network intrusion detection using a comprehensive dataset like UNSW-NB15. The developed models, including Random Forest, MLP, and LSTM, demonstrated promising performance in distinguishing between normal and malicious network activities. The utilization of Bayesian Gaussian Mixture Models for anomaly detection showcased an alternative approach, especially with the exploration of optimal thresholds through cross-validation. However, certain limitations must be acknowledged. The effectiveness of the models heavily relies on the dataset's representativeness, and biases or anomalies within the data may impact generalizability. Additionally, the dynamic nature of cyber threats poses challenges for static models, necessitating continuous adaptation and updates to stay relevant. This research advances the knowledge in the field of cybersecurity by demonstrating a practical application of AI that goes beyond traditional static defense mechanisms. By developing a model that learns from data and adapts to new threats, this study provides a framework for future AI systems in cybersecurity that are more resilient and capable of dealing with the dynamic nature of cyber threats. This contributes to a body of knowledge that supports the development of smarter, more adaptive NIDS, offering a significant improvement over the existing methods. For future work, there are several avenues for improvement and exploration. Firstly, incorporating more diverse and real-time datasets could enhance the models' ability to adapt to evolving cyber threats. Feature engineering and selection methodologies should be further optimized to extract more meaningful information from network traffic data.

Moreover, the research could benefit from the integration of explainability techniques to enhance the interpretability of model decisions. This would contribute to building trust in the model's predictions and facilitating the identification of false positives or negatives. The exploration of ensemble models or hybrid architectures, combining the strengths of different algorithms, could further enhance overall detection accuracy. Investigating the impact of hyperparameter tuning on model performance may also yield valuable insights into optimizing model configurations. In conclusion, this research lays the groundwork for future endeavors in the field of network intrusion detection. By addressing the mentioned limitations and embracing emerging technologies and methodologies, the potential for more robust and effective intrusion detection systems becomes a promising avenue for exploration.

A. Author Contributions

Conceptualization, M.A.; methodology, M.A.; software, M.A.; validation, M.A; formal analysis, M.A. and A.B.; investigation, M.A.; resources, M.A.; writing—original draft preparation, M.A.; writing—review and editing, S.M.; visualization, M.A.; supervision, S.M.; funding acquisition, M.A. All authors have read and agreed to the published version of the manuscript.

B. Funding

This research received no external funding.

C. Institutional Review Board Statement

Not applicable.

D. Informed Consent Statement

Not applicable.

REFERENCES

[1] Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling, and research directions. SN Computer Science, 2, 1-18.

[2] Ansari, M. F., Dash, B., Sharma, P., & Yathiraju, N. (2022). The Impact and Limitations of Artificial Intelligence in Cybersecurity: A Literature Review. International Journal of Advanced Research in Computer and Communication Engineering.

[3] Tao, F., Akhtar, M. S., & Jiayuan, Z. (2021). The future of artificial intelligence in cybersecurity: A comprehensive survey. EAI Endorsed Transactions on Creative Technologies, 8(28), e3-e3.

[4] Moustafa, Nour, and Jill Slay. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)." Military Communications and Information Systems Conference (MilCIS), 2015. IEEE, 2015.

[5] Das, R., & Sandhane, R. (2021, July). Artificial intelligence in cyber security. In Journal of Physics: Conference Series (Vol. 1964, No. 4, p. 042072). IOP Publishing.

[6] Sarker, I.H., Furhad, M.H. and Nowrozy, R., 2021. Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Computer Science, 2(3), p.173..

[7] Ansari, M.F., Dash, B., Sharma, P. and Yathiraju, N., 2022. The impact and limitations of artificial intelligence in cybersecurity: a literature review. International Journal of Advanced Research in Computer and Communication Engineering..

[8] Zaman, S., Alhazmi, K., Aseeri, M.A., Ahmed, M.R., Khan, R.T., Kaiser, M.S. and Mahmud, M., 2021. Security threats and artificial intelligence based countermeasures for internet of things networks: a comprehensive survey. Ieee Access, 9, pp.94668-94690.

[9] Kaur, R., Gabrijelčič, D. and Klobučar, T., 2023. Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, p.101804.

[10] Ansari, M. F., Sharma, P. K., & Dash, B. (2022). Prevention of phishing attacks using AI-based Cybersecurity Awareness Training. Prevention.

[11] Chomiak-Orsa, I., Rot, A. and Blaicke, B., 2019, August. Artificial intelligence in cybersecurity: the use of AI along the cyber kill chain. In International Conference on Computational Collective Intelligence (pp. 406-416). Cham: Springer International Publishing.

[12] Arora, J. B., & MH, L. Technologies used in Services and Impact of Those in Providing Security.

[13] Soni, V. D. (2020). Challenges and Solution for Artificial Intelligence in Cybersecurity of the USA. Available at SSRN 3624487.

[14] Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy. IEEE Access.

[15] Chan, L., Morgan, I., Simon, H., Alshabanat, F., Ober, D., Gentry, J., ... & Cao, R. (2019, June). Survey of AI in cybersecurity for information technology management. In 2019 IEEE technology & engineering management conference (TEMSCON) (pp. 1-8). IEEE.

[16] Stevens, T. (2020). Knowledge in the grey zone: AI and cybersecurity. Digital War, 1, 164-170.

[17] Basnet, A. (2022). Artificial intelligence in cyber security.

[18] Calderon, R. (2019). The benefits of artificial intelligence in cybersecurity.

[19] Khamzina, B., Roza, N., Zhussupbekova, G., Shaizhanova, K., Aten, A., & Meirkhanovna, B. A. (2022). Determination of Cyber Security Issues and Awareness Training for University Students. International Journal of Emerging Technologies in Learning (Online), 17(18), 177.

[20] Wiafe, I., Koranteng, F. N., Obeng, E. N., Assyne, N., Wiafe, A., & Gulliver, S. R. (2020). Artificial intelligence for cybersecurity: a systematic mapping of literature. IEEE Access, 8, 146598-146612.

[21] Wiafe, I., Koranteng, F. N., Obeng, E. N., Assyne, N., Wiafe, A., & Gulliver, S. R. (2020). Artificial intelligence for cybersecurity: a systematic mapping of literature. IEEE Access, 8, 146598-146612.

[22] Mohammed, I. A. (2020). Artificial intelligence for cybersecurity: A systematic mapping of literature. Artif. Intell, 7(9), 1-5.

[23] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[24] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[25] M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

Author 1 Mohammed Khalid Almuqrin

As a Master's student deeply immersed in the realm of cybersecurity, I'm dedicated to exploring the synergy between artificial intelligence and cyber defense. With a strong background in the industry, I bring practical insights and academic rigor to my research. My paper delves into the innovative applications of AI in fortifying cybersecurity measures, aiming to contribute valuable insights to the field.

Author 2 Prof. Shailendra Mishra

As a Professor in college of computer science , my expertise extends beyond cybersecurity into various other fields. With a diverse background, I've mentored students across disciplines, fostering innovative research and scholarly endeavors.

E-mail: 441104797@s.mu.edu.sa