American Journal of Technology Advancement 2997-9382, Vol 1, No. 1, 2024

Article

AI-Driven Cybersecurity for Business: Predictive Detection of Malware and Network Intrusion

Abdul Rahman Mohammad
Bachelor in Technology
Jawaharlal Nehru Technological University, Hyderabad

Mahbub Hasan
BSc in Information Technology
University of Information Technology and Sciences (UITS), Dhaka, Bangladesh

Jafrin Reza
Master of Science in Business Analytics
Trine University, USA

Mohd Salman Hossain Bhuiyan
Bachelor of Science in Computer Science and Engineering
East West University, Bangladesh

Abstract: As businesses grow ever more reliant on digital technologies, interdependent networks and information systems, malware attacks, unauthorized access and network intrusions have grown more dangerous. The traditional security mechanisms might not be effective in dealing with the changes in the threats, especially if malicious activities are not following the patterns they are familiar with or if they are attempting to beat the security mechanisms with non-traditional approaches. Artificial intelligence (AI) specifically machine learning can be used to improve cyber security by detecting patterns in network traffic and system memory to determine if it is malicious or not. The application of AI driven techniques in business for predictive malware detection and network intrusion detection is discussed. The study is based on two publicly available cybersecurity datasets: CIC-MalMem-2022, which provides support for detecting malware using memory analysis approaches; and CICIoT2023, which includes both malicious and benign network traffic data for various intrusion classes. Random Forest and XGBoost are suggested as the baseline machine-learning algorithms to assess both detection tasks. The datasets will be analysed separately since they are used to solve different cybersecurity problems and are composed of different sets of features. The evaluation of the model performance will be done based on the accuracy, precision, recall, F1 score, false positive rate and the model efficiency in terms of computation, with special emphasis on class imbalance and detection reliability. The study will also examine the role of AI-based detection capabilities in a business' risk management, incident response, digital asset protection, and possible impact on operations of a cyber-attack. Furthermore, implementation issues such as the risk of false alarms, computational needs, model generalization and data quality will be analyzed. In this study, the researchers used a comprehensive research platform to examine malware and network intrusion detection, in order to offer evidence-based findings regarding the possibilities and challenges of AI in cybersecurity for companies. The results of the experiments will be used to make recommendations for embedding machine-learning detection models into an organization's security plan and to highlight that simply being able to classify existing records doesn't mean they can be able to predict future attacks before malicious acts

happen.

Keywords: Artificial Intelligence, Cybersecurity, Malware Detection, Network Intrusion Detection, Machine Learning and Business Risk Management.



This is an open-access article under the CC-BY 4.0 license

1. Introduction

Background of the Study

Digital technologies have significantly changed how businesses operate, allowing them to operate through interconnected information systems to manage, provide services, communicate with stakeholders and transact with them. While digital transformation boosts efficiency and competitiveness, it simultaneously raises the risk of cybersecurity threats to the organization, such as malware attacks, unauthorized access, data breaches, and network intrusions [1]. All of these can have a negative impact on business continuity, on sensitive information, on financial losses and on customer confidence. Businesses are increasingly seeking cybersecurity plans that will allow them to promptly identify threats and manage risks properly. Most of the traditional cybersecurity methods that are used are based on pre-defined signatures, known attack patterns and security rules [2]. These mechanisms are still very useful, but may not be effective at recognizing new variants of malware, maliciously obfuscated applications, and network attacks. Threat detection can be enhanced by using AI, including machine learning algorithms, to detect and recognize patterns in system characteristics and network traffic that are correlated with malicious activity [3]. Using labeled datasets, algorithms like Random Forest and XGBoost might be able to classify observations as benign or malicious, in order to make cybersecurity analysis more systematic. In this research paper, two complementary datasets are explored to understand the use of AI in cybersecurity. CIC-MalMem-2022 facilitates malware detection based on memory analysis features like process information, dynamic-link libraries and system handles. CICIoT2023 is able to assist with the network intrusion detection, both by providing labeled traffic records of benign activities and several categories of cyberattacks [4]. These are the data sets that can be used for individual experiments on malware and network intrusion detection. Both datasets could be used to assess the success and failure of machine-learning techniques in various cybersecurity areas. This study is oriented by the performance of the classification and its implications for business security management, considering that it is necessary to be able to detect the previous occurrence of malicious actions successfully does not imply the capability to predict future attacks before they take place.

AI-Driven Predictive Detection of Malware and Network Intrusion

Predictive detection using AI means leveraging computational models to look for patterns that are linked to potential cybersecurity threats and to categorize activities that are observed based on risk characteristics [5]. Machine-learning algorithms can be used in malware detection to identify malicious activities compared to normal system activities in memory-related features. Obfuscation is a key technique used by malware that aims to hide the attacks or even bypass mainstream security systems, where memory analysis plays a significant role. The CIC-MalMem-2022 is a collection of memory images that includes both benign and malicious ones, with each malware category having its own set of benign and malicious images. These records can then be used to train models, which can then be used to assess the ability of various algorithms to classify the memory activity as malicious or benign [6]. Network intrusion detection is another, related, but different type of cybersecurity task. It includes analyzing the pattern of network traffic to discover unusual

communication patterns and classify possible attacks. The CICIoT2023 dataset covers network traffic from benign and various types of intrusion attacks, such as Distributed Denial-of-Service (DDS), Denial-of-Service (DoS), and Reconnaissance, Spoofing and Brute Force attacks. These records can be leveraged to assess machine-learning algorithms to detect malicious traffic and to differentiate attack types. The two data sets should be analyzed and evaluated separately, as the features of memory analysis are different to the features of network traffic. The baseline algorithms are proposed to be Random Forest and XGBoost for both the detection tasks. They will be evaluated based on precision, recall, F1-score, rate of false-positive and other suitable metrics [7]. Class imbalance (wider distribution of a few attack classes in comparison to others) will be addressed in particular. Computational efficiency and practical effects on business operations of detection errors will also be taken into consideration [8]. While such experiments can illustrate the effectiveness of models at classification of labeled observations, assertions of the ability to predict future attacks must be evaluated in future temporal or prospective experiments [8]. This separation is crucial when considering the real-world impact of AI on cybersecurity for businesses.

Problem Statement

Malware attacks and network intrusion poses are growing threats for businesses, but accurately detecting a wide range of threats is difficult. Familiar threats might not be detected by traditional security systems and AI-driven ones can generate false alarms, miss minority attack types, or fail to work as well on new data. In addition, the data characteristics of malware memory analysis and network intrusion detection are different, so it needs to be evaluated in a different way. Thus, the main research question of this study is the efficiency of application of the machine-learning models in determining malicious activities in the two cybersecurity tasks and how the efficiency of the machine-learning models can relate to the business security decisions. This study compares the performance of Random Forest and XGBoost on the CIC-MalMem-2022 and CICIoT2023 respectively, to evaluate the effectiveness, the reliability and the limitations of the two models.

Objectives of the study

The main objective of this research is to assess the potential of artificial intelligence for predictive malware detection and network intrusion detection systems in a business environment. The objectives of the study:

- To explore how well machine-learning algorithms can classify benign and malicious memory-analysis records with the CIC-MalMem-2022 dataset.
- To assess the effectiveness of the AI based models in identifying and categorizing network intrusions from CICIoT2023 dataset.
- To make the comparison of the classification performance of the Random Forest model and the XGBoost model in both detection tasks [9].
- To evaluate the performance of the model based on accuracy, precision, recall, F1-score, false-positive rate as per the classification task.
- To study the impact of class imbalance and data preprocessing on the accuracy of the malware and network intrusion detection.
- To explore how a secure cybersecurity model can be implemented with machine learning and its limitations.
- To research and formulate recommendations based on evidence for the incorporation of AI powered threat detection into business cyber security and risk management strategies.

Research Questions

The following research questions are used in guiding the study:

- What is the accuracy of classification of CIC-MalMem-2022 with respect to both the Random Forest and XGBoost classifiers?
- What is the capability of machine-learning models for detecting the network intrusion and

classification in CICIoT2023?

- What is the comparative performance of the two models (Random Forest and XGBoost) in terms of detection performance and false-positive rates in both datasets and computational efficiency?
- What are the main technical and organizational challenges associated with the embedding of artificial intelligence to detect and prevent malware and network intrusions in companies?

Significance of this study

This study is significant. The importance of this study lies in the examination of the possibility that Artificial Intelligence can help business cybersecurity with two complementary Business Cybersecurity applications Malware Analysis and Network Intrusion Detection. This allows the study to give empirical feedback on the ability and limitation of the machine-learning models to various cybersecurity issues by comparing CIC-MalMem-2022 and CICIoT2023 individually [10]. Significance for businesses. The study could provide insights to business managers and cybersecurity experts on the potential of AI-powered detection to aid in the detection of malicious activity, enhance security monitoring, and guide incident-response strategies [11]. The results can also help organizations assess how machine-learning tools can be integrated into their existing cybersecurity systems and what their benefits and constraints are. Importance to cybersecurity practitioners. Random Forest and XGBoost can be compared and this comparison can give some evidence about the classification performance, false positives, and computational requirements [12]. An analysis of the class imbalance and minority attack classes can also illustrate the need to consider overall accuracy alone as an inadequacy if one is using it to pick detection models. Significance for researchers. This study offers a framework that can be used for investigating the issue of malware detection and network intrusion detection with publicly available data sets and provides a comparative framework [13]. It also emphasizes the need to assess different cybersecurity tasks individually and critically assess the results of experiments in the light of the data. Business significance for business risk management. The study can help guide decisions about threat monitoring, resources allocation and investment in cybersecurity based on the performance of technical detections in relation to the security needs of the organization [14]. This study seeks to be a basis for a more evidence-based approach to the understanding of how detection using Artificial Intelligence can enhance business cybersecurity and the importance of human involvement, the regular assessment of models' performance and further testing of new threats.

Artificial Intelligence and Machine Learning in Cybersecurity

Artificial intelligence and machine learning are one of the key research areas in cybersecurity because they allow for automatic analysis of system activity, network traffic, and other digital artifacts related to malicious activity [15]. While traditional cybersecurity systems rely on known signatures and rules, machine-learning systems learn the patterns from past observations and use them to classify potentially harmful behavior [16]. Although they have been shown to be effective in certain applications, such methods may be used as tools for threat identification, anomaly detection and security monitoring, and their effectiveness relies on the quality and representativeness of the training data. In cybersecurity classification problems with structured data, Random Forest and XGBoost are of special interest. Random Forest is an ensemble of multiple decision trees, and XGBoost is an ensemble of sequentially improved decision trees using gradient boosting [17]. Both algorithms are able to model complex relationships between the input features and can serve as useful baseline algorithms for comparison of detection performance. Studies based on CIC-MalMem-2022 have included the analysis of traditional machine-learning methods as well as of more sophisticated models, in both cases highlighting the relevance of the comparative analysis of various algorithms when the experiments are conducted within the same set of conditions. But a high classification accuracy will not always ensure reliable cybersecurity protection [18]. Asymmetric datasets can lead to models overfitting the common classes and false-positive notifications can add

to the burden on security teams. Additionally, models can behave differently on unseen data than on the data that it was trained on. Thus, precision, recall, F1-score, false-positive rate and also computational efficiency should be assessed rather than solely accuracy when measuring cybersecurity research [19]. Explainability is also a critical consideration as business security professionals want to understand what is driving model predictions when exploring suspicious activity. In this current research, Random Forest and XGBoost serve as a foundation for assessing AI-based detection for two cybersecurity tasks [20]. They will be judged individually in CIC-MalMem-2022 and CICIoT2023. This method will enable the determination of strengths and weaknesses of machine-learning classification and explain the implications of the evidence that can come from it to inform business cybersecurity decisions.

AI-based Malware Detection using Memory Analysis

Malware detection is one of the most important challenges in cyber security as malware can harm confidential data, disrupt business operations and permit unauthorized access to organizational systems. A small number of malwares use "obscuration" methods to hide their behavior and evade traditional methods of detection [21]. The characteristics of processes, dynamic-link libraries, system handles, and other features of memory can also be used to provide evidence by performing memory analysis. These can be used as features to differentiate benign system activity from patterns related to malicious software by machine-learning models. The aim of CIC-MalMem-2022 is to facilitate memory analysis research for obfuscated malware detection. It contains malware samples for ransomware, spyware and Trojan horse as well as benign samples [22]. The data set thus serves as a yardstick to compare and assess whether machine-learning algorithms can detect malicious patterns in structured memory features [23]. This data set has been analyzed using various classification methods, such as Random Forest, Gradient Boosting and Deep Learning. The results of these studies show that the performance of the models should be considered in light of the preprocessing techniques, feature selection, class definitions, and evaluation methods [24]. While all the above are opportunities, there are still a few methodological challenges in the realm of memory-based malware detection. There can also be redundant features, class distributions can impact how reliable predictions are and models may have been trained on specific malware samples and not generalize to future threats [25]. Very high accuracy must be taken with a pinch of salt and the evaluation procedure should be checked to ensure that there is no data leakage and an appropriate separation between training and testing observations. Furthermore, successful classification of the memory records doesn't necessarily imply that the detection performance will be successful in an operational business environment [26]. In the current study, these two models Random Forest and XGBoost will be evaluated for the task of benign and malicious memory-analysis records using CIC-MalMem-2022. Multiple measures will be used to assess performance such as precision, recall, F1 score, and false positive rate. Results will offer insights into the effectiveness and the challenges of AI-powered malware detection in the realm of business cybersecurity.

AI-Based Network Intrusion Detection and Business Cybersecurity

An intrusion detection system (IDS) is a network security tool that detects attempts to intrude into a network. As businesses rely on increasingly interconnected systems and devices from the Internet of Things, the security of these networks is vital in protecting information, ensuring service availability, and supporting operational continuity. Denial of service attacks, distributed denial of service attacks, reconnaissance, spoofing, brute force and botnet activities are examples of network attacks [27]. An intrusion detection system based on machine learning can be used to examine the features of network flows and classify them as normal or malicious, and if necessary, differentiate among attack types. Neto et al. (2023) presented CICIoT2023 as an IoT environment study dataset and benchmark to detect large-scale attacks [28]. The dataset was created with the help of a testbed that comprises different IoT devices and is composed of benign traffic and various types of malicious traffic. It includes structured network-flow records which are used for the assessment of binary and multiclass intrusion detection using machine-learning techniques [29]. The dataset is significant for the current research due to its uniformity, which allows the study of the ability of AI models to

differentiate between normal traffic and various types of malicious network activity. The issues with network intrusion detection models include the problems of class imbalance, changing attack behavior and false-positive alerts, as well as computational complexity [30]. A model which works well on a random split test set may not work the same way if it is deployed in another network environment or tested against attacks that it has not seen before. The restriction is significant for companies since it may enable threats to go undetected or lead to security people receiving useless alerts [31]. The research of intrusion-detection has therefore focused on methods that incorporate the traditional machine-learning techniques with anomaly detection to enhance the ability of identifying unfamiliar attacks. The current experiment will employ CICIoT2023 to test the performance of two classifiers such as XGBoost and Random Forest in network intrusion detection [32]. The evaluation will feature binary classification as well as multiclass attack classification, if possible, due to the computing power involved [33]. The accuracy, precision, recall, false positive rate, and computational efficiency will be used as metrics to assess the results. This study will evaluate the potential of AI-driven network monitoring for supporting cybersecurity decisions, taking into account these factors, along with business security needs.

Existing Research and Research Gap

Manoharan and Sarker (2022) discussed the possibilities of Artificial Intelligence (AI) and Machine Learning (ML) in the field of cyber security and next generation threat detection. This article covers topics of anomaly detection, behavioral pattern analysis, predictive threat identification, natural language processing, deep learning, neural networks and automated threat detection and response. The authors also recognize problems of data privacy and ethics. The article does not present an exhaustive experimental comparison of predictive models based on a limited set of malware or network-intrusion samples, however, but rather offers a general overview of AI applications [2]. The information found in the articles is not enough to draw valid conclusions about the performance of the models under the same experimental conditions and to conduct detailed evaluations based on the results of precision, recall, F1-score and false-positive rate. Thus, there is a demand for future research to examine the impact of dataset imbalance, machine-learning algorithms for malware classification and network intrusion detection, and the impact of detection performance on business cybersecurity and incident response, as well as organizational resilience.

Raji et al. (2023) explored how artificial intelligence (AI), machine learning (ML), and data analytics can be used to enhance cybersecurity threat detection and response. Their research demonstrates that such technologies are able to deal with vast amounts of security information, uncover previously unknown patterns, detect new threats, and contribute to automatic response to cyberattacks. The authors also highlight the significance of data analytics for enhancing security analysis, and for allowing cybersecurity solutions to become dynamic based on the information provided by the attacks. The study used a literature based approach with a search using various electronic databases such as IEEE Xplore, ACM Digital Library, ScienceDirect, Scopus and Google Scholar. The research, however, does not clearly state the experimental comparison of predictive models for Malware and Network Intrusion Detection, as the abstract seems more of a holistic discussion on integrated cybersecurity technologies [3]. Additional studies are required to analyze the performance of machine learning algorithms with the real dataset, compare the performance of machine learning algorithms with the same metric, and study the efficiency of the algorithms for the classification of various kinds of malware and network attacks. Solving these problems could offer real-world examples for the use of AI-powered predictive detection systems in business cybersecurity and organizational risk management.

Das et al. (2022) discussed the threats and opportunities of artificial intelligence (AI)-based cybersecurity intrusion detection. They have explored the potential of AI and machine learning solutions for combating cybercrime, identifying cyberattacks, and enhancing data security in an organization. The authors point to the ability of AI-driven systems to detect new attacks sooner and help develop proactive defenses to them [4]. The review also takes into account the preparations made by industries to defend against the changing cyber threat. The article, however, was more

concerned with providing a general overview of AI-driven cybersecurity applications than detailing the experimental assessment of predictive models with malware and network-intrusion datasets. There is still a lot more work to be done to evaluate the performance of various machine learning algorithms with the same evaluation metrics, assess their performance on various types of malware attacks and network attacks and understand the impact of imbalanced data sets. Closing these gaps can facilitate the understanding of the role predictive detection models may play in business cyber security, enhancing the identification and understanding of threats and bolstering organizational resilience.

2. Methodology

The research methodology used in this study is quantitative and experimental, which will be used to investigate AI-based cybersecurity for business, specifically in the areas of predictive malware detection and network intrusion detection [34]. The methodology aims to assess machine-learning models on two publicly available cybersecurity datasets: CIC-MalMem-2022 and CICIoT2023. Malware detection is analyzed using memory-analysis capabilities with CIC-MalMem-2022 and malicious network traffic is analyzed using CICIoT2023. The data is analyzed individually since the feature structure of each data set is different and they include a variety of cybersecurity tasks [35]. The methodology proposed is data collection, data preprocessing, exploratory data analysis, feature examination, development of the model, and evaluation of the model performance [36]. The Support Vector Machine (SVM) model is also added to the malware classification experiment, and if the experiments are carried out, the Random Forest and XGBoost models can be considered as other base models. The study utilizes quantitative performance metrics to evaluate classification effectiveness and how the findings might be relevant to business cybersecurity and risk-management decisions [37]. A focus on data quality, class imbalance, data leakage, and the difference between threat classification and forecasting future attacks.

Research Design

Research is quantitative, experimental and supervised machine learning [38]. This method is suitable as the study seeks to evaluate the accuracy of AI algorithms at detecting benign activity from malicious ones from labeled cybersecurity data. There are two independent experiments: malware detection and network intrusion detection [39]. To classify benign and malicious memory-analysis records, malware experiment is conducted with the CIC-MalMem-2022 dataset, while network experiment is conducted with the CICIoT2023 dataset, which is used to understand benign and attack traffic [40]. The experimental design, focusing on repeatability and consistent process and testing with proper metrics. Before developing a model, descriptive analysis is done to learn more about class distributions, the relationships between features, and potential data quality problems [41]. This leads to an interpretation of the evidence within the framework of business cybersecurity and not on the grounds that a good score on a set of data equates to a good score in operational environments.

Data Collection and Dataset Description

The study relies on secondary data, which is publicly available cybersecurity data created for research purposes. The data sets contain labelled observations and features that are structured for supervised classification. CIC-MalMem-2022: The total number of memory-analysis records is 58,596, consisting of 29,298 benign memory-analysis and 29,298 malware memory-analysis records. It protects against ransomware, spyware and Trojan horses [42]. The dataset includes memory characteristics, such as process activity, dynamic-link libraries, system handles, and so on. It is used to research how to classify benign and malicious memory behavior. CICIoT2023: This dataset includes those attacks that can be used for research on IoT network intrusion. The sample of traffic records uploaded into the Excel file for the preliminary analysis is of 1000 records of which 26 are benign and 974 are attack traffic [43]. Attacks are included in the sample with attack labels of DDoS, DoS and others. However, the results obtained from this uploaded file are very small and highly imbalanced and should not be interpreted as representing the entire CICIoT2023 data set without

further confirmation. Data are maintained separately throughout the research as their features, labels and research goals vary.

Data Preprocessing

Data preprocessing is done to pre-process the datasets for machine-learning analysis [44]. The first steps involve looking at the column names, finding the variables to target, checking data types, looking for missing data, and looking for infinite or invalid numerical observations. Inappropriate variables are not included in the models by separating numerical features from identifier and target-label columns [45]. The missing numerical values could be replaced with training-set medians and observations that are duplicated are examined to see if they are a true repeat or if a potential data-leakage scenario exists [46]. The target variable for CIC-MalMem-2022 is to determine whether the record is benign or a malware record. The stratified sampling for the experiment on the classification of the malware contains 12,000 observations: 6,000 observations are sampled from each of the classes [47]. The sample is split into 9000 observations for training and 3000 observations for testing, with 75:25 split. The stratified split ensures the same class balance in the training and testing subsets. The target labels are explored for the purpose of CICIoT2023, which is to classify them as benign and attack, and if possible, as multiclass attack [48]. There is a very strong class imbalance in the uploaded sample, which should be taken into account when choosing evaluation techniques. To prevent information leakage, any imputation, scaling, feature selection or resampling is fitted based on the training data.

Exploratory Data Analysis

Before model evaluation, both datasets are analyzed using exploratory data analysis in order to look at the structure and the characteristics of both datasets. Bar charts show the distribution of benign and malicious records; category-distribution charts display the distribution of various families of malware and network attack labels [49]. Box plots are used to compare selected numerical features between observations of benign and malicious. Correlation heatmaps are used to determine linear relationships among memory-analysis and network-traffic features. In this analysis of CIC-MalMem-2022, the process count, malware categories, and relationships between selected memory features are analyzed [50]. The analysis of CICIoT2023 covers the following aspects: traffic-label frequencies, benign/attack distribution, flow-rate characteristics, and network-feature correlations [51]. These visualizations can be used to detect potential feature redundancy and class imbalance [52]. But descriptive patterns are not enough to prove that a feature can be used to capture the malicious behavior, this is a matter of model evaluation.

Machine-Learning Model Development

The malware classification experiment is based on the Support Vector Machine (SVM) machine learning algorithm having a radial basis function kernel. The numerical features are scaled as they can also impact SVM performance if they have different scales [53]. The missing values are imputed using median, and the model is trained on the training subset and then tested with the held-out test subset. The receiver operating characteristic (ROC) curve and area under the curve (AUC) are calculated using the predicted class probabilities [54]. The test accuracy and ROC AUC for the initial SVM experiment were 99.90% and 1.000 respectively. The results are considered preliminary as other investigations are required to exclude leakage due to duplication, leakage due to features, and the sample-specific effect [55]. The random forest and XGBoost models are considered as candidates for comparison. But, the result of their efforts should be reported only after the models have been trained and evaluated. The network intrusion experiment should also be designed and tested independently with CICIoT2023.

Model Evaluation Metrics

To accuracy the study considers several measures of classification accuracy.

Metric	Purpose
Accuracy	Measures the proportion of correctly classified observations
Precision	Measures the proportion of predicted positive observations that are actually positive
Recall	Measures the proportion of actual positive observation correctly identified
F1-Score	Combines precision and recall through their harmonic mean.
False-positive rate	Measures the proportion of negative observations incorrectly classified as positive
ROC-AUC	Measures how effectively a model separates the two classes across classification thresholds

It is used to determine true positives, true negatives, false positives and false negatives, and this is achieved by the use of a confusion matrix [56]. ROC curves illustrate the relationship between the true-positive rate and false-positive rate at different thresholds. In this study, since the uploaded CICIoT2023 sample is highly imbalanced, it is crucial to look at the precision, recall, F1 score and class wise performance to understand the performance of this CICIoT2023 sample.

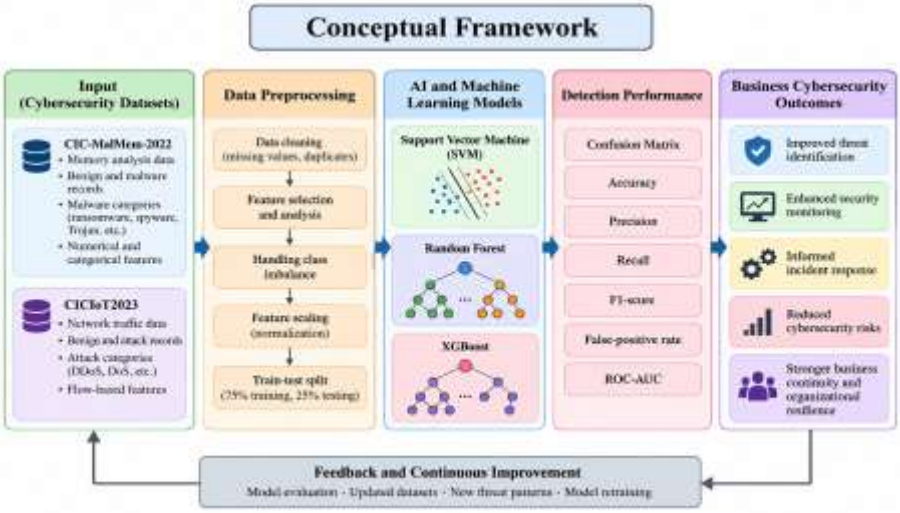
Research Reliability and Validity

Documentation of datasets and their preprocessing, sampling strategy, model configuration, and evaluation methods help ensure reliability [57]. Where applicable, the fixed random seed is used to ensure the same results are obtained for reproduction [58]. The separation of training and testing observations allows to estimate performance on test observations not directly used for model fitting [59]. A few extra measures need to be taken for validity. Apparent performance can be due to duplicate observations, correlated samples, target leakage and inappropriate feature selection [60]. Potential for reduplication, independent testing and splits by time and/or group (if appropriate and the structure of the data allows). Classification accuracy does not necessarily mean effectiveness in the real world or predict future attacks before they happen using these data sets.

Ethical Considerations

This study is based upon public and free cyber security data sets that are used for defensive research and analysis [61]. The investigation is more about detecting malicious activity, comprehending model performance and making business security decisions. All datasets should be attributed, licensed and the origin should be acknowledged [1]. Results should be presented transparently, with any limitations of the methods used and potential sources of bias. No argument or promise of 'effective operation' can be made without proper testing in representative environments.

Conceptual Framework



This flowchart depicted integrates cybersecurity datasets, preprocessing, machine learning model, detection evaluation and business risk management for continuous improvement

The conceptual framework as shown in this research called "AI-Driven Cybersecurity for Business: Predictive Detection of Malware and Network Intrusion" explains the relationship between the cybersecurity datasets, data preprocessing, machine-learning algorithms, threat-detection performance, and business cybersecurity outcomes [2]. It suggests the quality of the input data and the choice of suitable machine-learning models affect the effectiveness of malware and network intrusion detection [3]. It contains two datasets namely CIC-MalMem-2022, which is a dataset for the detection of malwares, and CICIoT2023, which is a dataset for the detection of network intrusions [4]. Each of these datasets is composed of cybersecurity tasks and thus processed independently with appropriate preprocessing and classification steps [5]. Data quality, feature characteristics, class distribution and model selection are all important analytical aspects of the framework [6]. There are several machine-learning algorithms which can be used to evaluate the performance of classification, such as Random Forest, XGBoost and Support Vector Machine (SVM). The accuracy, precision, recall, F1-score, false-positive rate, and ROC-AUC are used to evaluate the resulting predictions, as applicable. The measures are evidence about the effectiveness of the models to detect benign activity and malicious activity [7]. The last stage links the detection findings to potential business outcomes, such as better security monitoring, more informed incident response, decreased cybersecurity risk exposure, and better organizational resilience [8]. These are potential as opposed to guaranteed effects of model performance.

3. Results and Analysis

The findings obtained from this research on AI-Driven Cybersecurity for Business: Predictive Detection of Malware and Network Intrusion are presented in this section. This analysis relies on two cybersecurity datasets CIC-MalMem-2022 and CICIoT2023 for studying the characteristics of malware and network intrusion patterns [9]. The results are presented as descriptive data analysis, feature correlation analysis and machine-learning model evaluation. The first analysis examines the distribution of the benign and malicious records, different types of malware, and characteristics of memory processes, network traffic classes and correlations between selected features [10]. These observations indicate the nature of the datasets and give an indication to potential problems related to the classification using machine learning [11]. Support Vector Machine (SVM) model is tested using a confusion matrix and receiver operating characteristic (ROC) curve to determine its accuracy on identifying malware records from benign records [12]. The results are analyzed with reference to the level of detection reliability and implications for business cybersecurity applications, but validation in real environments is still needed to ensure generalizability of the results.

Distribution of Benign and Malware Records

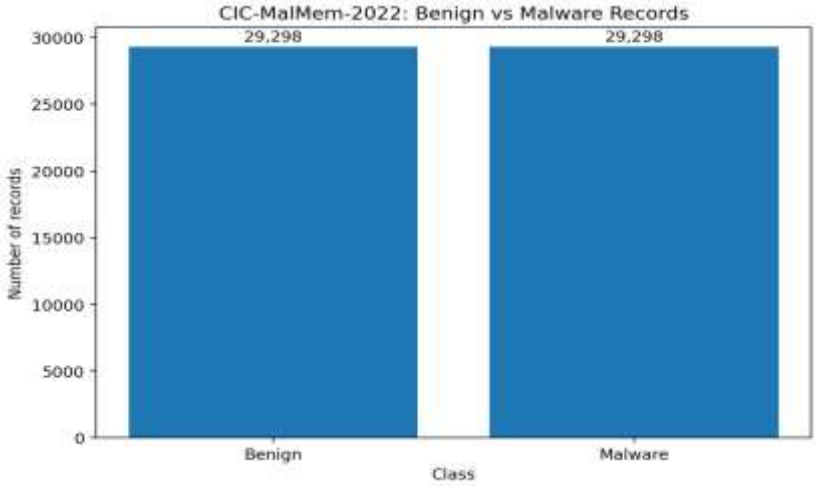


Figure 1. This image shows the distribution of all types of records in CIC-MalMem-2022 are equal

Class distribution of the CIC-MalMem-2022 data set used in this study is shown on the bar chart. The data set has 58,596 data points, with 29,298 benign data points and 29,298 malware data points [13]. The classes are balanced, with 50% of the data being in each class. This equal distribution is helpful to assess machine learning algorithms as neither of the classes is overwhelming the data set. The data set will thus serve as an appropriate basis for comparing the models such as Random Forest and XGBoost, with respect to the ability to distinguish between the benign and malicious. But, the model performance still needs to be achieved by training and testing.

Distribution of Malware Categories

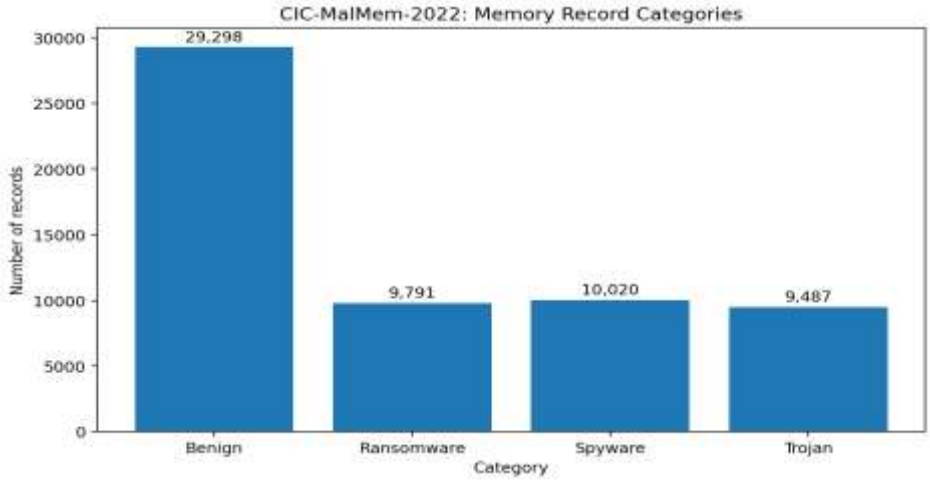


Figure 2. This image shows the distribution of benign records and three categories of malware

The figure below shows a bar chart of the distribution of benign records and three types of malware in the CIC-MalMem-2022 dataset. The greatest number of observations are for benign records (29,298). Spyware has the highest number of records with 10,020, followed by ransomware with 9,791 records and Trojan horses with 9,487 records [14]. The distribution of the three categories of malware is fairly uniform and offers chances to compare the characteristics of the malware in terms of memory analysis [15]. The distribution suggests that the data set extends beyond one type of malware but encompasses various types [16]. The diversity enables further analysis of machine-learning models for malware classification and analysis of potential use of the models in business cybersecurity and threat detection.

Correlation Analysis of Malware Memory Features

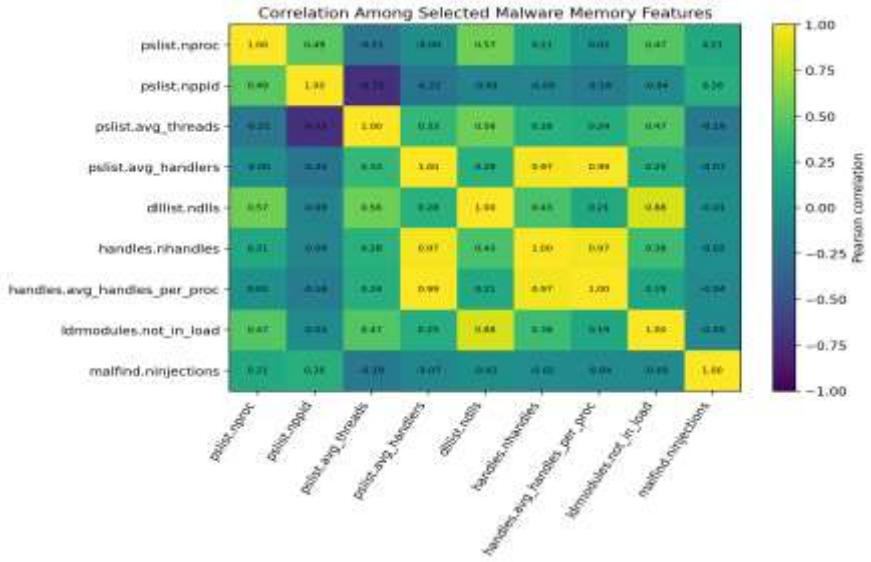


Figure 3. This image shows the correlations between the selected features of memory analysis for AI-based malware detection

The heatmap shows the Pearson correlation values between the nine chosen features of the CIC-MalMem-2022 dataset for memory-analysis [17]. The positive correlation between pslist.avg_handlers and handles.avg_handles_per_proc is strong (0.99) and handles. Handles and handles.avg_handles_per_proc is strong (0.97). Also relatively high (0.88) is the correlation between dlllist.ndlls and ldrmodules.not_in_load. On the other hand, pslist.avg_threads has a high negative correlation with pslist.nppid of −0.73. The relationships suggest that certain memory characteristics are correlated, and others are antithetical [18]. The knowledge of feature correlations can be used to aid feature selection and to identify redundancy in creating machine-learning models for malware detection.

Comparison of Process Counts for benign and Malware records

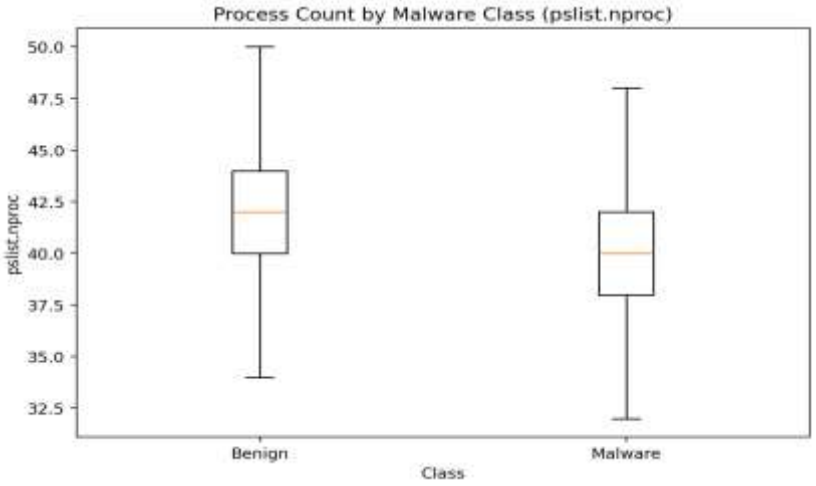


Figure 4. This image shows a comparison of the distributions of process-counts in benign and malware memory records

Box plot of the distribution of the process count (pslist.nproc) for benign and malware records in the CIC-MalMem-2022 data set. The median count of processes in benign records is around 42, while it's 40 for malware records [19]. The interquartile range for benign records is around 40 to 44 while that of malware records is 38 to 42. The whiskers show that benign process counts range between about 34 and 50, whereas the counts for malware are between 32 and 48. The differences

indicated here seem to be fairly small between the two classes in terms of process-count distributions [20]. The significant commonality, however, suggests that process count might not be a useful metric for differentiating a malicious application from normal traffic [21]. This could be useful when used in combination with other memory properties to enhance machine-learning classification [22].

Distribution of Benign and Attack Network Traffic

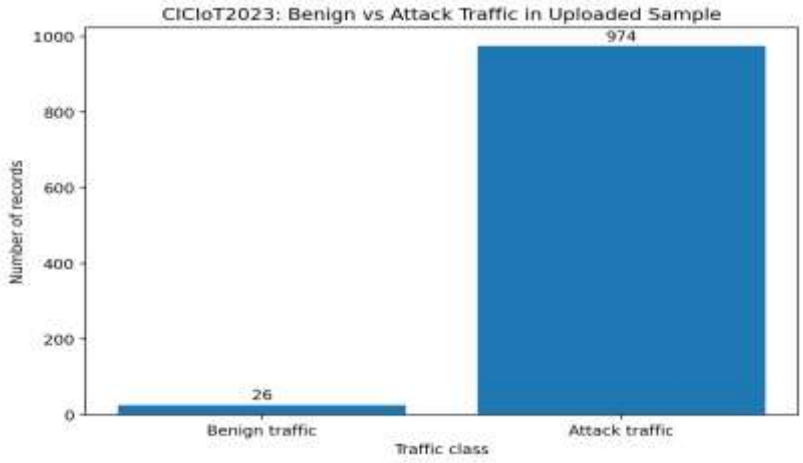


Figure 5. This image shows the imbalance, with benign traffic at 57% and Attack traffic at 43%, of the traffic

The bar chart displays the distribution of the benign and malicious network traffic for the uploaded sample of the CICIoT2023. There are 1000 records in the sample, 26 benign traffic records and 974 attack traffic records. 97.4% of the sample is from attackers while only 2.6% is benign traffic [23]. The large gap is a clear sign of a disproportionately large number of observations of malicious traffic being uploaded. The dataset contains examples that can be used for analyzing network intrusion detection, but the small amount of benign examples in the dataset may impact model training and evaluation [24]. Thus the use of stratified sampling and class-imbalance handling techniques and class-specific performance metrics should be considered [25]. These measures can be used to assess whether there is a successful discrimination between normal activities and attacks by machine-learning models or whether they are biased towards the majority class [26].

Correlation Analysis of Network Traffic Features



Figure 6. This image illustrates relations between selected features of network traffic for AI based intrusion detection

The heatmap displays the Pearson correlation coefficients between a specific set of network traffic features of the network traffic data provided in the CICIoT2023 sample. The highest positive correlation was found to be between Rate and State with a value of 1.00, which shows a near linear

variation [27]. There is a moderately strong positive correlation between Tot size and Variance (0.66) and Duration and Variance (0.42). Also, there is a positive correlation between the two variables, Protocol Type and Tot size with coefficient of correlation 0.36. However, the IAT data has a negative correlation with flow_duration of −0.23. Relationships can be used to detect potentially redundant features and to inform feature selection [28]. But correlation doesn't equal causation and it doesn't prove which of the features is best for intrusion detection.

Comparison of network flow rates for benign and attack traffic

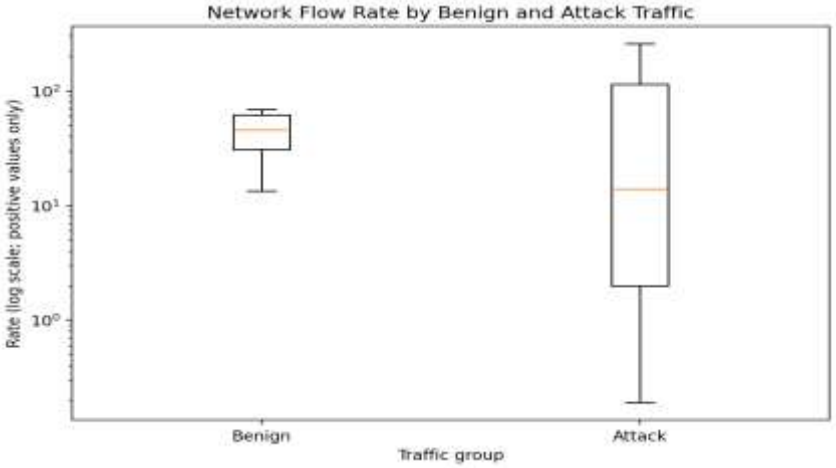


Figure 7. The image illustrates the difference in the distribution of network flow-rate between benign traffic and attack traffic

The box plot illustrates the comparison of network flow rates of benign traffic vs attack traffic in the uploaded CICIoT2023 sample. The median rate for the benign group is about 45 whereas the median rate for the attack group is about 14. The attack group is also much more varied, ranging from less than 1 on the logarithmic scale to more than 200. Benign traffic, on the other hand, has a smaller spread with the majority of samples falling within the range of about 30 to 60. These differences indicate that these flow-rate characteristics could potentially be used to differentiate some types of attacks from normal network traffic [29]. There is significant intra attack variation, however, and flow rate is not necessarily a good indicator of all attacks [30]. This feature could be used in conjunction with other network parameters to enhance the machine-learning detection performance.

Distribution of the Most Frequent Network Attack Categories

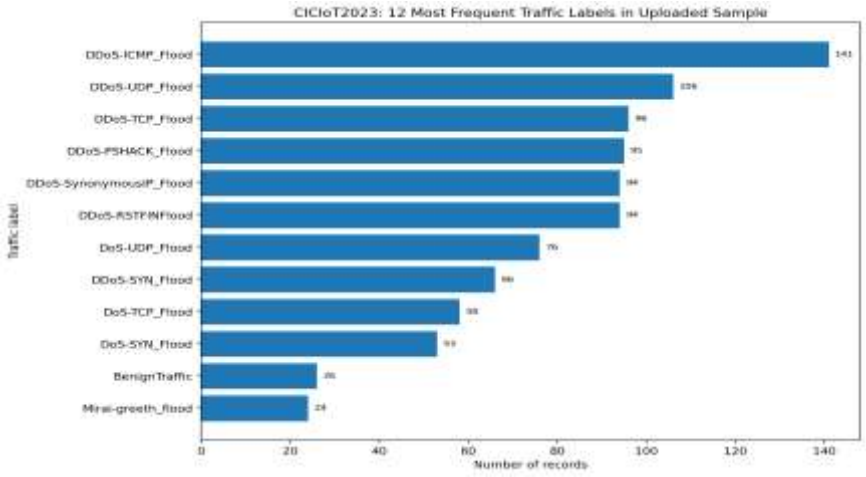


Figure 8. This image illustrates the most frequent network traffic labels, highlighting the dominance of DDoS attacks

The horizontal bar chart displays the 12 most common traffic labels found in the CICIoT2023 sample uploaded. DDoS-ICMP_Flood is the most common one, 141 records; DDoS-UDP_Flood with 106 records and DDoS-TCP_Flood with 96 records. The number of records in DDoS-PSHACK_Flood is equal to 95, in DDoS-SynonymousIP_Flood - equal to 94 and in DDoS-RSTFINFlood - equal to 94. DoS-UDP_Flood is responsible for 76 records and DDoS-SYN_Flood for 66 records. BenignTraffic has 26 records and Mirai-greeth_flood has 24 records. These results show that the attack with the DDoS feature is the most prevalent attack on the uploaded sample [31]. This distribution emphasizes the need to consider the performance of machine-learning models across various categories of attacks, as well as the overall classification performance [32].

Performance Evaluation of the SVM Malware Classification Model

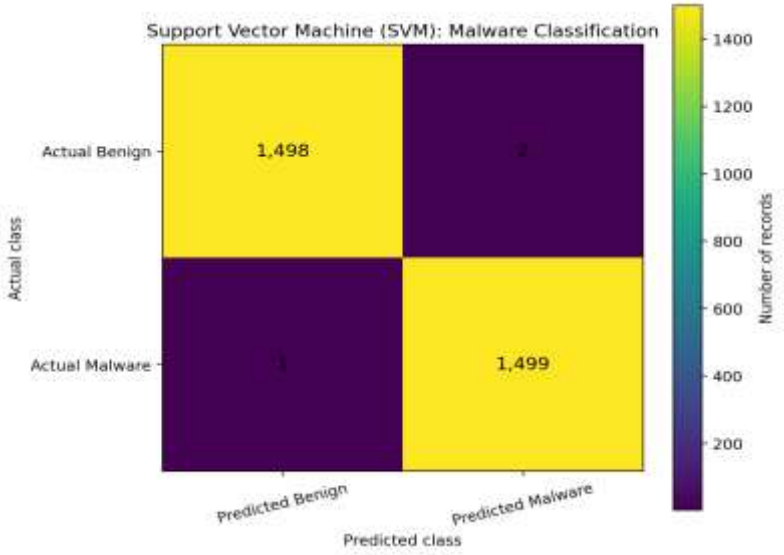


Figure 9. This confusion matrix illustrates SVM performance in distinguishing benign and malware records

The confusion matrix shows the accuracy of the Support Vector Machine (SVM) model when classifying benign and malicious records from the CIC-MalMem-2022 dataset. 1,498 benign records were correctly classified and 1,499 malware records were correctly classified. It incorrectly labeled two benign records as malware, giving it false-positive predictions, and one record containing malware as benign, which is a false negative [33]. The model was found to be accurate with 3000 test observations with an accuracy of around 99.90%. The results demonstrate good classification accuracy on the test sample selected. This performance has not been extensively tested to confirm it works when faced with malware that has not yet been seen, and in business settings [34]. Before making definite conclusions, it is essential to take into consideration the potential for data leakage, duplication of records and a sample selection effect.

ROC Curve Analysis of the SVM Malware Detection Model

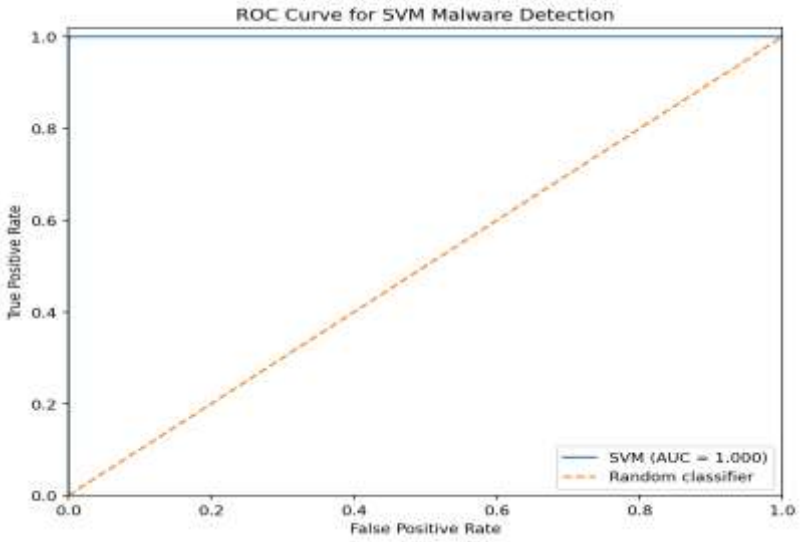


Figure 10. This image demonstrates a ROC curve for the SVM model to classify the benign records from malware

The receiver operating characteristic (ROC) curve is used to assess the performance of the Support Vector Machine (SVM) model to classify benign records as non-malicious and malware as malicious in the CIC-MalMem-2022 dataset. The curve is near the upper left corner, suggesting that there may be a high true positive rate with a low false positive rate over the range of classification thresholds assessed [34]. The area under the curve (AUC) of the model is 1.000 indicating the model discrimination is good on the selected test sample [35]. The dashed diagonal line is the performance of a random classifier, and serves as a baseline for comparison [36]. The AUC observed demonstrates good classification performance, but the maximum value should be taken with a grain of salt. Before drawing conclusions about the model's performance in the real world of business cybersecurity, it would be beneficial to validate further, using tests for duplicate data, data leakage, and generalization to independent test data.

4. Discussion and Analysis

This study's results show the potential of Artificial Intelligence (AI) as a tool for supporting cybersecurity of businesses by detecting malware and analyzing network intrusions. The CIC-MalMem-2022 dataset comprises 58,596 records, with half being benign and the other half being malware. This balanced distribution serves as a useful basis to assess machine-learning based classification models before a first set of imbalances between the two classes [37]. Spyware is the largest category of malware with 10,020 records; ransomware comes in second place with 9,791 records; and Trojans are in third place with 9,487 records. These are the diverse classes of malicious software contained in the dataset, and exhibit the potential to be explored as differences in memory-analysis characteristics [38]. A correlation analysis also shows high correlations between some of the memory characteristics, such as the average process handlers with the average handles per process (0.99), average process handlers with the total number of handles (0.97). These relationships can suggest feature redundancy which should be taken into account when designing a successful malware detection system and feature selection and model evaluation [39]. The process-count comparison also shows differences between the benign and malware records, with a median of ~42 and ~40 respectively. The significant overlap in their distributions suggests that the number of processes might not be enough to give a good discrimination, which reinforces the idea to use more than one process [40]. The analysis of CICIoT2023 gives more insights into the nature of network intrusion [41]. There are 1000 records in the uploaded sample - 974 attack and 26 benign traffic records. This distribution is highly skewed, and may affect the outcome of machine-learning algorithms and give inaccurate estimates of accuracy unless remediated properly. Attack-label

analysis reveals that DDoS-ICMP_Flood has the highest number of records (141), followed by DDoS-UDP_Flood (106) and DDoS-TCP_Flood (96). Based on these observations, it appears that the sample that is uploaded focuses on DoS activities, and is not balanced for the various categories of network activities [42]. The network feature correlation analysis shows that there is a perfect correlation of 1.00 between Rate and Srate, and a positive correlation of 0.66 between Tot size and Variance [43]. Such relationships can be used to identify redundant predictors and for feature selection. The flow-rate box plot reveals more variation in the attack records than in benign records, suggesting that when other characteristics of the network are added, flow-rate characteristics could be used to help with intrusion detection [48]. The results of SVM malware classification experiment shows that 1,498 benign records and 1,499 malware records were classified correctly in the test sample, which gives an accuracy of 99.90%. Two FP/One FN errors were found in the confusion matrix and the AUC of the ROC curve was 1.000. While these results may appear excellent at discriminating in the sample tested, they should be used with caution as the results may be due to duplicate records, feature leakage, or patterns in the sample [49]. The results highlight the analytical capabilities of AI tools in cybersecurity and the necessity of validation by independent sources, careful preprocessing of data, proper evaluation criteria and testing on representative data prior to using these models in real business security scenarios.

Future Work

The future research direction for the study, inclusion of more diverse data sets and the testing of the feasibility of applying AI in business settings. While the present study has focused on CIC-MalMem-2022 and CICIoT2023, additional cybersecurity datasets, such as recent variants of malware, previously unseen network attacks and various enterprise traffic patterns could be introduced for future research [50]. This would help to test if the machine learning models can be generalized to data not seen in training and testing. Additionally, more algorithms such as Random Forest, XGBoost, Support Vector Machine, deep neural networks and ensemble learning algorithms should be evaluated to find the best models for the purpose of the detection with low computation cost and operational needs [51]. Another crucial direction is to create temporal prediction models that can detect emerging attack patterns early at the time where significant damage is still not done [52]. To determine if a network is being accurately predicted, and to distinguish it from the classification of already observed malicious activity, a series of time ordered observations was necessary, along with clearly-defined forecasting targets and suitable evaluation procedure. In the future, the class imbalance problem could be solved by adequate sampling techniques, cost-sensitive learning algorithms, and class-weighting, especially for benign traffic or the specific types of attacks that are underrepresented [53]. Also, feature selection and dimensionality reduction methods can be explored to ensure that the computation of the features is kept simple but still captures meaningful metrics of malicious behavior [54]. Another promising research area is Ex explainable intelligence: Cybersecurity experts and business managers are always in need of understandable explanations on model predictions when exploring incident investigation and security decisions [55]. Explainability techniques, for instance, could be used in future studies to pinpoint the most significant characteristics of memory and network-flows to classify threats [56]. Experiments involving network traffic streaming, automated alerts generation and integration with security information and event management systems should also be considered in real-time implementation [57]. These investigations might include assessing detection latency, processing needs, frequency of false alarms, and the capability of sustaining performance under varying network circumstances [58]. The adversarial robustness, concept drift, privacy protection and the ethical aspects of automated security decision-making should be studied in future. Fusion security models that integrate artificial intelligence (AI) surveillance capabilities and traditional security policies, processes and controls along with human intervention may be beneficial for organizations [59]. Independent validation of the research should be done on different data sets, and with temporally separated data, carefully selected train-test splits, and external test environments to detect data leakage problems and evaluate the generalization ability of the models [60]. This research will build on this work, offering

more empirical evidence to support the benefits, scale, and usefulness of AI in cybersecurity [61]. These advances could enable companies to better detect threats, allocate cybersecurity resources more efficiently and enhance their organization's capabilities to withstand current and emerging malware and network intrusion attacks.

5. Conclusion

The results of this study could contribute to developing the potential of artificial intelligence and machine learning technologies to detect malware and suspicious network activities in business environments. The analysis of the CIC-MalMem-2022 and CICIoT2023 data sets gave insights into categories of malware, characteristics of the processes that were being targeted, characteristics of the networks being attacked, and relationships between important features. In the first Support Vector Machine (SVM) experiment, conducted on a stratified subsample of the CIC-MalMem-2022 data set, the accuracy rate was 99.90% and the ROC-AUC score was 1.000, which shows that the classification performance was good for the experimental conditions. These findings, however, must be confirmed with other data sets and intrusion traffic to ascertain if the same results can be obtained with independent data sets and intrusion traffic in real life. The exploratory analysis also highlighted the need to consider class distributions, feature relationships and dataset imbalances prior to building predictive models. The small number of benign records from the uploaded CICIoT2023 sample highlights the need for careful evaluation through multiple performance measures and not just accuracy. For reliable and practical performance, the precision, recall, F1-score, false-positive rate and computational needs should all be taken into account when designing effective predictive detection systems. Additionally, it is crucial to continuously monitor and update machine learning models to ensure they remain effective in addressing emerging threats, ensure security data is protected, and that there's human oversight. Further research is needed to compare them under the same experimental settings, to test them with previously unknown threats, and to consider other algorithms such as Random Forest and XGBoost. Further research is needed for the explain ability of models, computational efficiency and more importantly to change attack patterns. While the results are not conclusive proof of effectiveness in the real-world, they are a stepping stone towards further research on data-driven cybersecurity approaches. In conclusion, businesses can enhance their ability to detect threats, minimize security risks, and build operational resilience in today's digital landscape by integrating trusted predictive models with proven security controls, robust incident response strategies, and risk management practices.

REFERENCES

[1]. Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., & Ghorbani, A. A. (2023). CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. Sensors, 23(13), 5941.

[2]. Manoharan, A., & Sarker, M. (2023). Revolutionizing cybersecurity: Unleashing the power of artificial intelligence and machine learning for next-generation threat detection. DOI: https://www. doi. org/10.56726/IRJMETS32644, 1.

[3]. Raji, A., Olawore, A., Mustapha, A., & Joseph, J. (2023). Integrating Artificial Intelligence, machine learning, and data analytics in cybersecurity: A holistic approach to advanced threat detection and response. World Journal of Advanced Research and Reviews, 20(3), 2005-2024.

[4]. Dash, B., Ansari, M. F., Sharma, P., & Ali, A. (2022). Threats and opportunities with AI-based cyber security intrusion detection: a review. International Journal of Software Engineering & Applications (IJSEA), 13(5).

[5]. Jabed, M. M. I., Khawer, A. S., Ferdous, S., Niton, D. H., Gupta, A. B., & Hossain, M. S. (2023). Integrating Business Intelligence with AI-Driven Machine Learning for Next-Generation

Intrusion Detection Systems. International Journal of Research and Applied Innovations, 6(6), 9834-9849.

[6]. Sunkara, G. (2022). AI-driven cybersecurity: Advancing intelligent threat detection and adaptive network security in the era of sophisticated cyber attacks. Well Testing Journal, 31(1), 185-198.

[7]. Chirra, D. R. (2023). Towards an AI-Driven Automated Cybersecurity Incident Response System. International Journal of Advanced Engineering Technologies and Innovations, 1(01), 429-451.

[8]. Umar, H., & Abbas, A. (2022). AI-Powered Threat Intelligence: Enhancing Cybersecurity with Predictive Analytics and Machine Learning.

[9]. Yang, M. H. (2022). AI-driven cybersecurity: Intrusion detection using deep learning. Multidisciplinary Innovations & Research Analysis, 3(4), 1-14.

[10]. Pissanidis, D. L., & Demertzis, K. (2023). Integrating AI/ML in cybersecurity: An analysis of open XDR technology and its application in intrusion detection and system log management.

[11]. Pavan Kumar, P., Satish, M., Sunitha Devi, B., Prakash, A., Pradeep Reddy, K., & Malli Babu, S. (2023, December). The future of AI in predicting cybersecurity threats. In International Conference on [12]. Data Science, Machine Learning and Applications (pp. 1382-1395). Singapore: Springer Nature Singapore.

[13]. Gaioto, F. (2023). Big Data Intrusion Detection Using AI-Based Supervised Classifiers and Machine Learning Ensembles for Cybersecurity Threat Prevention. Unpublished Work.

[14]. Bodade, N. (2023). AI-Driven Intrusion Detection System for Zero-Day Attack Prevention. Available at SSRN 6992938.

[15]. Ibiyeye, O., Okolo, J. N., & Adeniji, S. A. (2022). A Comprehensive Evaluation of AI-Driven Data Science Models in Cybersecurity: Covering Intrusion Detection, Threat Analysis, Intelligent Automation, and Adaptive Decision-Making Systems. Communication In Physical Sciences, 8(4), 745-763.

[16]. Tarek, J. H., & Rahman, W. (2023). AI-driven cybersecurity, IoT networking, and resilience strategies for industrial control systems: A systematic review for US critical infrastructure protection. International Journal of Scientific Interdisciplinary Research, 4(4), 144-176.

[17]. Rajaram, S. (2022). AI-driven threat detection: Leveraging big data for advanced cybersecurity compliance. Educational Administration: Theory and Practice.

[18]. Rockey, H. (2022). AI-Driven Cybersecurity in IoT: Detecting and Preventing Attacks on Smart Devices. Journal of Cybersecurity and Artificial Intelligence, 3(2), 110-128.

[19]. Hassan, M. U. (2023). Study of artificial intelligence in cyber security and the emerging threat of AI-driven cyber attacks and challenge. Available at SSRN 4652028.

[20]. Rahman, M. K., Dalim, H. M., & Hossain, M. S. (2023). AI-Powered solutions for enhancing national cybersecurity: predictive analytics and threat mitigation. International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence, 14(1), 1036-1069.

[21]. Farooq, U. (2023). Cyber-physical security: AI methods for malware/cyber-attacks detection on embedded/IoT applications (Doctoral dissertation, Politecnico di Torino).

[22]Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., & Pospelova, V. (2022). The emerging threat of ai-driven cyber attacks: A review. Applied Artificial Intelligence, 36(1), 2037254.

[23]. Sultanan, S., Rahman, M. M., Hossain, M. S., Gony, M. N., & Rafy, A. L. (2022). AI-powered threat detection in modern cybersecurity systems: Enhancing real-time response in enterprise environments. World Journal of Advanced Engineering Technology and Sciences, 6(2), 136.

[24]. Hassan, M. M., Abrar, M. F., & Hasan, M. (2023). An explainable AI-driven machine learning

framework for cybersecurity anomaly detection. Cyber security and business intelligence, 197-219.

[25]. Guntupalli, R. (2023). AI-driven threat detection and mitigation in cloud infrastructure: Enhancing security through machine learning and anomaly detection. Available at SSRN 5329158.

[26]. Aramide, O. O. (2023). Predictive analytics and automated threat hunting: The next frontier in AI-powered cyber defense. International Journal of Technology, Management and Humanities, 9(04), 72-93.

[27]. Wolsey, A. (2022). The state-of-the-art in AI-based malware detection techniques: a review. arXiv preprint arXiv:2210.11239.

[28]. Ariyo, S. M. I. E. E. (2023). The future of cyber security in an AI-driven world. The Cable.

[29]. Saxena, D., Gupta, I., Gupta, R., Singh, A. K., & Wen, X. (2023). An AI-driven VM threat prediction model for multi-risks analysis-based cloud cybersecurity. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 53(11), 6815-6827.

[30]. Akram, E., & Basit, F. (2023). AI-powered information security: Innovations in cyber defense for cloud and network infrastructure.

[31]. Zahid, W., & Jack, W. (2023). Cybersecurity for the Future: How AI Enhances Threat Prediction and Response.

[32]. Adebayo, Y. (2023). Intelligent AI Driven Cloud Native Security Framework for Enterprise Systems Financial Platforms IoT Networks and Real Time Threat Detection. International Journal of Computer Technology and Electronics Communication, 6(5), 7345-7353.

[33]. Saurabh, K. (2023). Optimizing it Program Management in the ERA of AI-driven Cybersecurity Solutions. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 15(03), 391-396.

[34]. Gopalsamy, M. (2023). AI-Driven Solutions for Detecting and Mitigating Cyber Threats on Social Media Networks. Int. J. Adv. Res. Sci. Commun. Technol, 3(2).

[35]. Karunamurthy, A., Kiruthivasan, R., & Gauthamkrishna, S. (2023). Human-in-the-loop intelligence: Advancing AI-centric cybersecurity for the future. Quing: International Journal of Multidisciplinary Scientific Research and Development, 2(3), 20-43.

[36]. Ismail, B. I., Abdul, S., Khan, S. M., Sattar, S. A., & Muhammad, S. (2023). AI for cyber security: Automated incident response systems. Available at SSRN 5477114.

[37]. Maddireddy, B. R., & Maddireddy, B. R. (2023). Enhancing network security through AI-powered automated incident response systems. International Journal of Advanced Engineering Technologies and Innovations, 1(02), 282-304.

[38]. Banerjee, S., & Parisa, S. K. (2023). Ai-enhanced intrusion detection systems for retail cloud networks: A comparative analysis. Transactions on Recent Developments in Artificial Intelligence and Machine Learning, 15, 15.

[39]. Wang, J. (2023). AI/ML-Powered Cybersecurity and Cloud Computing Strategies for Optimized Business Intelligence in ERP Cloud. Res. Prepr.

[40]. Aslam, S., & Jackson, M. (2022). AI-Driven Anomaly Detection: Strengthening Data Protection in Enterprise Networks.

[41]. Andersson, M. (2023). AI-Powered Enterprise Cloud Frameworks for Secure Data Integration and Predictive Threat Intelligence. International Journal of Engineering & Extended Technologies Research (IJEETR), 5(5), 7282-7290.

[42]. Samuel, B. (2023). AI-Driven Cyber Threat Detection and Response. International Journal of Intelligent Automation & Robotics Engineering, 6(1), 01-07.

[43]. Rahman, M. M., HOSSAIN, M. S., Mashfiquer, M., Rahman, M. S. U., NAHAR, S., & Rahman,

M. M. (2023). Reinforcement learning for adaptive cybersecurity: AI-driven threat detection and response mechanisms. ICONIC Research and Engineering Journals, 7(1), 721-732.

[44]. Vasa, Y., & Singirikonda, P. (2022). Proactive cyber threat hunting with AI: Predictive and preventive strategies. International Journal of Computer Science and Mechatronics, 8(3), 30-36.

[45]. Adeyemi, D. S. (2023). Autonomous response systems in cybersecurity: A systematic review of AI-driven automation tools. Communication in Physical Sciences, 9(4), 878-898.

[46]. Mohamed, N. (2023). Current trends in AI and ML for cybersecurity: A state-of-the-art survey. Cogent Engineering, 10(2), 2272358.

[47]. Nicki, P. (2022). Real-Time Threat Intelligence: AI-Based Approaches to Cyber Risk Prediction and Mitigation.

[48]. Ali, M. A., & Alqaraghuli, A. (2023). A Survey on the Significance of Artificial intelligence (AI) in Network cybersecurity. Babylonian Journal of Networking, 2023, 21-29.

[49]. Adebayo, A. S., Chukwurah, N., & Ajayi, O. O. (2022). Proactive ransomware defense frameworks using predictive analytics and early detection systems for modern enterprises. Journal of Information Security and Applications, 18(2), 45-58.

[50]. Soundappan, S. J. (2023). Generative AI-Driven Intelligent Cybersecurity Framework for Secure Hybrid Cloud Enterprise Systems. International Journal of Emerging Trends in Engineering and Management Research, 8(6), 14728.

[51]. Ayo, A., George, C. K., & Smart, E. (2023). Overview of AI-powered cybersecurity and threat intelligence.

[52]. Maka, S. R., Jha, K. M., Chinta, P. C. R., Moore, C. S., Katnapally, N., & Sadaram, G. (2023). AI-powered big data and ERP systems for autonomous detection of cybersecurity vulnerabilities. Nanotechnology Perceptions, 1, 46-64.

[53]. Smith, S. (2023). AI-driven cybersecurity: Leveraging big data for advanced threat detection and risk mitigation. Unpublished manuscript.

[54]. Bushigampala, B. K., & Inaganti, A. C. (2023). Threat detection in critical infrastructure using AI models. International Journal of Acta Informatica, 2(1), 196-208.

[55]. Ashraf, H., & Bhutto, B. (2022). Preventing Tomorrow's Threats Today: AI-Driven Innovations in Cyber Defense.

[56]. Singh, B., Anand, A., & Prabhat, S. (2022). Real-time network monitoring and incident response with AI-driven automation data center and WAN transformation. Available at SSRN 5331665.

[57]. Dari, S. S., Thool, K. U., Deshpande, Y. D., Aush, M. G., Patil, V. D., & Bendale, S. P. (2023). Neural Networks and Cyber Resilience: Deep Insights into AI Architectures for Robust Security Framework. Journal of Electrical Systems, 19(3), 78.

[58]. Akande, J. O., Raji, O. M. O., Babalola, O., Abdulkareem, A. O., Samson, A., & Folorunso, S. (2023). Explainable AI for Cybersecurity: Interpretable Intrusion Detection in Encrypted Traffic.

[59]. Tauseef, A. (2023). Database Technologies in AI: Transforming Cybersecurity with Automated Threat Detection Systems.

[60]. Kumar, A. (2023). Next Generation AI Driven Security and Analytics Framework for Cloud Native Enterprise Systems Financial Platforms and IoT Infrastructure. International Journal of Engineering & Extended Technologies Research (IJEETR), 5(6), 7689-7697.

[61]. Maddireddy, B. R., & Maddireddy, B. R. (2022). Cybersecurity threat landscape: Predictive modelling using advanced AI algorithms. International Journal of Advanced Engineering Technologies and Innovations, 1(2), 270-285.

[62]. Dataset Link https://www.kaggle.com/datasets/subhajournal/iotintrusion

https://www.kaggle.com/datasets/luccagodoy/obfuscated-malware-memory-2022-cic