UNIVERSITY OF MPUMALANGA

DICT431: RESEARCH METHODOLOGY

ASSIGNMENT 3 2024

PERIOD	**:**	**TERM 2**
MODULE DESCRIPTION	**:**	**RESEARCH METHODOLOGY**
MODULE CODE	**:**	**DICT431**
FACULTY	**:**	**AGRICULTURE & NATURAL SCIENCES**
QUALIFICATION	**:**	**ADVANCED DIPLOMA IN ICT**
DUE DATE	**:**	**29h May 2024**
EXAMINER	**:**	**Prof N Wayi-Mgwebi**

.

INSTRUCTIONS

- THE GENERAL UNIVERSITY OF MPUMALANGA POLICIES, PROCEDURES AND RULES PERTAINING TO WRITTEN ASSESSMENTS APPLY TO THIS ASSESSMENT.
- COPYING AND PLAGIARISM IS A SERIOUS OFFENSE, A MARK OF ZERO WILL BE AWARDED IF IDENTIFIED TO HAVE BEEN COPIED.
- SUBMIT YOUR ASSIGNMENT ON MOODLE FOR TURN-IT-IN CHECK
- CLEARLY REFERENCE YOUR SOURCES
- STUDENTS WHO SUBMITTED IDENTICAL WORK WILL BOTH BE ALLOCATED ZERO (0) MARK.
- PLEASE USE THE INDIVIDUAL COVER PAGE

TABLE OF CONTENTS

Contents

Research Topic:

Development of a Self-learning and Adaptive Machine Learning-based Intrusion Detection System for Real-time Cyber Threat Detection

1.1. Introduction / Background :

In recent years, cybersecurity has become a critical issue due to the growing frequency and complexity of cyber-attacks. These attacks have advanced significantly, presenting serious risks to individuals, organizations, and nations. Traditional intrusion detection systems (IDS), which rely primarily on signature-based techniques, have long been central to cybersecurity defenses. However, these methods are increasingly inadequate against sophisticated and novel threats (Jemili, Rahma, & Quajdi, 2023). Signature-based IDS can only detect known threats, leaving systems exposed to zero-day attacks and other advanced persistent threats (APTs) that do not match any known signatures (Sowmya & Mary, 2023).

The evolving nature of cyber threats calls for more advanced and adaptable solutions. In this regard, machine learning-based intrusion detection systems (ML-IDS) have gained significant attention. ML-IDS use machine learning algorithms to analyze large volumes of network data, identify patterns, and detect anomalies that may indicate malicious activity (Chauhan & Singh, 2020). Unlike traditional IDS, ML-IDS do not rely solely on predefined rules or signatures; they can learn from new data, adapt to evolving threats, and improve their detection capabilities over time.

Research into ML-IDS has shown their potential to enhance intrusion detection effectiveness. Techniques such as anomaly detection and deep learning allow these systems to identify subtle deviations from normal behavior, thereby recognizing previously unknown threats (Mishra et al., 2019). This capability to detect zero-day attacks and other sophisticated threats makes ML-IDS a promising solution for modern cybersecurity challenges.

The increasing sophistication of cyber-attacks underscores the need for robust and adaptive intrusion detection systems. Cybercriminals continually refine their tactics, techniques, and procedures to evade detection, making it essential for cybersecurity solutions to evolve accordingly (Jemili, Rahma, & Quajdi, 2023). Therefore, there is an urgent need for innovative IDS architectures that can effectively combat the ever-changing threat landscape and protect digital assets.

1.2. Statement of the Problem :

The rapid evolution and growing complexity of cyber threats present significant challenges for organizations aiming to safeguard their digital assets. Traditional intrusion detection systems (IDS), which mainly use signature-based detection methods, are increasingly proving insufficient. These systems can only identify threats they already know about by comparing incoming data to a database of known signatures. This limitation leaves them unable to detect new or zero-day attacks (Buczak & Guven, 2016). As cyber attackers constantly develop new techniques to evade existing security measures, the gap between traditional IDS capabilities and the evolving threat landscape continues to widen.

The shortcomings of traditional IDS are especially problematic given the increasing frequency and sophistication of cyber-attacks. Advanced Persistent Threats (APTs), sophisticated malware, and other types of cyber-attacks often use obfuscation techniques to avoid detection by signature-based IDS. As a result, organizations remain vulnerable to data breaches, financial losses, and reputational harm (Singh & Silakari, 2015).

The inability of traditional IDS to adapt to new and unknown threats highlights the need for more advanced and adaptable intrusion detection systems. Machine learning-based intrusion detection systems (ML-IDS) have emerged as a promising solution. These systems use advanced algorithms to analyze network data, detect anomalies, and identify patterns that indicate malicious activity, offering the potential to uncover previously unknown threats (Dhanasekar & Ravichandran, 2019). However, for ML-IDS to be effectively implemented and perform well in real-world scenarios, extensive research and evaluation are necessary to address issues such as high false positive rates and the substantial computational resources required (Sommer & Paxson, 2010).

1.3. Research questions/ Hypothesis :

1.3.1. Research Questions:

i. How effective are machine learning-based intrusion detection systems (ML-IDS) in identifying and mitigating modern cyber threats in real-time compared to traditional intrusion detection systems?

ii. What are the key features and capabilities required for an ML-IDS to adapt to and effectively counter evolving cyber threats?

iii. How does the performance of ML-IDS vary across different types of cyber threats, including zero-day attacks, malware, and Advanced Persistent Threats (APTs)?

iv. What are the challenges and limitations associated with the implementation and deployment of ML-IDS in real-world organizational settings?

1.3.2. Hypothesis:

i. **Hypothesis 1:** Machine learning-based intrusion detection systems (ML-IDS) are more effective than traditional intrusion detection systems (IDS) in detecting and mitigating sophisticated and novel cyber threats in real-time.

ii. **Hypothesis 2:** The adaptability and learning capabilities of ML-IDS significantly enhance their performance in identifying unknown and zero-day cyber threats compared to signature-based IDS.

iii. **Hypothesis 3:** The integration of advanced machine learning algorithms, such as deep learning and anomaly detection, in ML-IDS improves the system's ability to detect a wide range of cyber threats, including those that evade traditional detection methods.

1.4. Research Objectives :

1.4.1. Primary Objective:

- The primary objective will be to evaluate the efficacy of machine learning-based intrusion detection systems (ML-IDS) in real-world cybersecurity scenarios, focusing on their ability to detect and mitigate modern cyber threats.

1.4.2. Secondary Objectives:

i. Investigate the performance of ML-IDS compared to traditional intrusion detection systems (IDS) in detecting various types of cyber threats, including zero-day attacks, malware, and Advanced Persistent Threats (APTs).

ii. Assess the adaptability of ML-IDS to evolving cyber threats and their capacity to learn from new data, thereby enhancing detection capabilities over time.

iii. Analyze the key features and capabilities of ML-IDS that contribute to their effectiveness in detecting and responding to cyber threats, including the integration of advanced machine learning algorithms and anomaly detection techniques.

iv. Explore the challenges and limitations associated with the implementation and deployment of ML-IDS in organizational settings, considering factors such as scalability, resource requirements, and integration with existing security infrastructure.

v. Propose strategies for optimizing the deployment and operation of ML-IDS in real-world cybersecurity environments, aiming to improve overall cybersecurity posture and resilience against cyber threats.

1.5. Significance of the study :

This study holds significant importance as it has the potential to enhance cybersecurity measures and counteract the escalating threats posed by cyber-attacks. Through evaluating the effectiveness of machine learning-based intrusion detection systems (ML-IDS) in real-world cybersecurity scenarios, valuable insights into advanced intrusion detection methods are offered. The findings of this research carry implications for various aspects of cybersecurity strategy. Firstly, organizations can utilize the comparison between ML-IDS and traditional systems to strengthen their cybersecurity posture, thereby reducing the susceptibility to successful cyber-attacks (Mansfield-Devine, 2019). Secondly, by assessing the adaptability of ML-IDS to evolving threats, this study aids in the development of proactive cybersecurity strategies, enabling organizations to stay ahead of emerging dangers (European Union Agency for Cybersecurity, 2020). Furthermore, by identifying the key features of ML-IDS that enhance their effectiveness, incident response protocols can be refined, ensuring swift detection and mitigation of cyber threats (National Institute of Standards and Technology, 2018). Lastly, by recognizing the challenges associated with the implementation and deployment of ML-IDS, organizations can optimize resource allocation, thereby maximizing the efficacy of their cybersecurity investments (Cybersecurity and Infrastructure Security Agency, 2021).

2. Literature review :

2.1. Evolution of Cyber Threat Landscape:

In recent years, the cybersecurity landscape has undergone notable changes, marked by a surge in the frequency and intricacy of cyber threats (Jemili, Rahma, & Quajdi, 2023). Traditional intrusion detection methods, like signature-based techniques, have faced challenges in adapting to these swift advancements, leaving organizations susceptible to sophisticated and emerging cyber-attacks (Sowmya & Mary, 2023).

2.2. Rise of Machine Learning-Based Intrusion Detection Systems:

To address the shortcomings of traditional intrusion detection methods, there has been a rising interest in machine learning-based intrusion detection systems (ML-IDS) (Chauhan & Singh, 2020). ML-IDS harness sophisticated machine learning algorithms to scrutinize network data, pinpoint anomalies, and recognize patterns suggestive of malicious behavior. Unlike conventional IDS, ML-IDS have the capacity to adjust to emerging threats and glean insights from fresh data, thus enhancing their detection prowess over time (Mishra et al., 2019).

2.3. Effectiveness of ML-IDS:

Research studies have demonstrated the effectiveness of ML-IDS in enhancing cybersecurity measures and mitigating cyber threats. Ahmad et al. (2018) compared the performance of different machine learning algorithms, including Support Vector Machine, Random Forest, and Extreme Learning Machine, for intrusion detection purposes. Their study found that these algorithms exhibited varying degrees of effectiveness in detecting cyber threats, with Support Vector Machine outperforming the other algorithms in terms of detection accuracy and false positive rate.

2.4. Challenges and Limitations:

Even though ML-IDS have promise, they also encounter some hurdles. Zhang and Lee (2000) pointed out a problem with scalability, saying ML-IDS might have trouble handling big loads of network data effectively. Also, Kim et al. (2014) mentioned that ML-IDS could be vulnerable to tricky maneuvers used by clever attackers, which makes people worry about how dependable they are in real-world cybersecurity situations.

2.5. Future Directions:

Moving forward, there will be a need for further research to address the challenges and limitations associated with ML-IDS and to explore new approaches for enhancing their effectiveness. Javaid et al. (2016) proposed a deep learning approach for network intrusion detection, which has shown promising results in improving detection accuracy and reducing false positive rates. By continuing to innovate and refine machine learning techniques for intrusion detection, researchers can contribute to the development of more robust and reliable cybersecurity solutions.

3.1. Research Paradigm, Methodology, and Approach :

3.1.1. Research Paradigm:

This study adopts a pragmatic research paradigm, prioritizing practical solutions to real-world problems. Pragmatism allows for flexible methods to address research questions, emphasizing the practical application of findings and blending qualitative and quantitative data (Creswell & Plano Clark, 2018). In assessing ML-IDS, pragmatism advocates for adaptable methodologies to address evolving cyber threats and organizational contexts.

3.1.2. Research Methodology:

The study employs a mixed methods approach, integrating qualitative and quantitative methods for a comprehensive analysis of ML-IDS.

- **Quantitative Methods**: Empirical evaluation of ML-IDS performance using statistical analysis, including detection rate, false positive rate, and computational efficiency. Data collection from simulated and real-world network environments ensures robust findings.

- **Qualitative Methods**: Expert interviews, case studies, and thematic analysis provide insights into challenges and benefits of ML-IDS implementation in real-world settings, enhancing understanding and informing recommendations.

3.1.3. Quantitative Approach:

i. **Data Collection**: Simulation of network environments to generate data on cyber threats and use of real-world datasets from cybersecurity repositories.

ii. **Data Analysis**: Evaluation of ML-IDS performance using metrics such as detection accuracy and comparative analysis with traditional IDS.

3.1.4. Qualitative Approach:

i. **Interviews**: Semi-structured interviews with cybersecurity experts to gather insights on ML-IDS implementation challenges.

ii. **Case Studies**: Development of case studies to explore real-world application, challenges, and benefits of ML-IDS.

iii. **Document Analysis**: Review of relevant documentation to supplement findings from interviews and case studies.

3.2. Population and Sampling :

3.2.1. Population:

The population for this study consists of:

i. **Cybersecurity Systems**: This includes a range of intrusion detection systems (IDS), both traditional and machine learning-based (ML-IDS), used in various sectors such as corporate environments, educational institutions, healthcare facilities, and governmental organizations. These systems are chosen to represent different deployment scenarios and threat landscapes.

ii. **Cybersecurity Professionals**: Individuals with expertise in cybersecurity, including IT security managers, network administrators, cybersecurity analysts, and researchers. These professionals are actively involved in the deployment, management, and evaluation of IDS and have practical experience with the challenges and benefits associated with these systems.

3.2.2. Sampling for Cybersecurity Systems:

i. Sample Size:

- A total of 10-15 IDS, including both traditional and ML-IDS, will be selected for evaluation. This sample size is considered adequate to provide a representative understanding of the performance and adaptability of different systems across various settings.

ii. Sampling Technique:

- **Purposive Sampling**: This technique will be employed to select IDS that are representative of different types of environments and deployment scales. Systems will be chosen based on specific criteria such as the type of machine learning algorithms used, the scale of deployment, and the availability of performance data. This method ensures that the selected systems are relevant to the study's objectives.

3.2.3. Sampling for Cybersecurity Professionals:

i. Sample Size:

- Approximately 20-30 cybersecurity professionals will be interviewed. This number is sufficient to achieve saturation in qualitative data and to ensure a diverse range of perspectives on the implementation and effectiveness of ML-IDS.

ii. Sampling Technique:

- **Snowball Sampling**: Initially, a few key cybersecurity experts will be identified through professional networks and industry contacts. These experts will then recommend additional professionals who have relevant experience and insights. This method helps in reaching a broader network of knowledgeable individuals.
- **Stratified Sampling**: To ensure balanced representation from different sectors (corporate, educational, healthcare, governmental), the sample will be stratified accordingly. This approach ensures that insights are gathered from a variety of contexts where IDS are deployed.

3.2.4. Data Collection Methods:

i. For Cybersecurity Systems:

- **Performance Data**: Collect data on detection accuracy, false positive rates, and computational efficiency from the selected IDS.
- **System Logs and Reports**: Analyze logs and reports generated by IDS to evaluate their performance and identify patterns in threat detection.

ii. For Cybersecurity Professionals:

- **Interviews**: Conduct semi-structured interviews to gather qualitative insights on the challenges, benefits, and practical considerations of implementing ML-IDS.

- **Questionnaires**: Use structured questionnaires to collect quantitative data on the perceived effectiveness and limitations of IDS from a larger sample of professionals.

3.2.5. Data Analysis

i. **Quantitative Data**: Statistical analysis will be conducted on performance metrics to compare traditional IDS and ML-IDS. Techniques such as t-tests or ANOVA may be used to determine significant differences in performance.

ii. **Qualitative Data**: Thematic analysis will be performed on interview transcripts and open-ended questionnaire responses to identify common themes and patterns. This analysis will provide deeper insights into the practical experiences of cybersecurity professionals.

3.3. Data Collection :

3.3.1. Data Collection for Cybersecurity Systems

The data collection process for this study involves gathering both quantitative and qualitative data to comprehensively evaluate the performance and implementation challenges of machine learning-based intrusion detection systems (ML-IDS).

3.3.2. Quantitative Data Collection

i. Simulated Network Environments:

  - **Setup**: Create simulated network environments using virtual machines and network simulation tools to replicate various cyber-attack scenarios, including zero-day attacks, malware, and Advanced Persistent Threats (APTs).
  - **Data Generation**: Generate traffic data that includes both benign and malicious activities. This simulated data will help in testing the detection capabilities of the ML-IDS under controlled conditions.
  - **Metrics Collection**: Collect data on key performance metrics such as detection rate, false positive rate, precision, recall, and computational efficiency.

ii. Real-World Data Sets:

- **Publicly Available Data Sets**: Use publicly available cybersecurity data sets from repositories like the UNSW-NB15 dataset, the CICIDS2017 dataset, and the KDD Cup 1999 dataset. These datasets contain labeled instances of normal and attack traffic.
- **Metrics Collection**: Like the simulated environments, collect performance metrics for the ML-IDS using these real-world data sets.

iii. System Logs and Reports:

- **Logs Analysis**: Analyze system logs and incident reports generated by IDS during their operation in real-world networks. These logs provide valuable information on the types of threats detected, the response times, and any missed detections.
- **Performance Metrics**: Extract relevant metrics from these logs to complement the data from simulated environments and public data sets.

3.3.3. Qualitative Data Collection

i. Interviews with Cybersecurity Professionals:

- **Participant Selection**: Select 20-30 cybersecurity professionals using snowball and stratified sampling techniques to ensure diverse representation across different sectors.
- **Semi-Structured Interviews**: Conduct semi-structured interviews to gather detailed insights into the practical challenges and benefits of implementing ML-IDS. Interview questions will focus on topics such as system deployment, operational challenges, integration with existing security infrastructure, and user experiences.
- **Recording and Transcription**: Record and transcribe interviews for subsequent analysis. This will ensure accurate capturing of participants' insights and facilitate thorough thematic analysis.

ii. Questionnaires:

- **Design**: Develop structured questionnaires with both closed and open-ended questions to collect quantitative data on the perceived effectiveness and limitations of IDS from a larger sample of professionals.
- **Distribution**: Distribute the questionnaires through online survey platforms and professional networks to reach a broader audience.
- **Analysis**: Analyze the responses using statistical methods for closed-ended questions and thematic analysis for open-ended questions.

iii. Case Studies:

- **Selection**: Identify and develop case studies of organizations that have implemented ML-IDS. Case studies will provide real-world examples of how these systems are used, the benefits observed, and the challenges encountered.
- **Data Collection**: Collect data through document analysis, direct observations (where possible), and interviews with key personnel involved in the deployment and management of the IDS.

3.3.4. Data Analysis

i. Quantitative Data Analysis:

- **Statistical Methods**: Use statistical techniques such as t-tests, ANOVA, and regression analysis to compare the performance of traditional IDS and ML-IDS. Evaluate detection accuracy, false positive rates, and computational efficiency.
- **Visualization**: Create visual representations such as graphs and charts to illustrate the performance metrics and facilitate easier comparison and interpretation of results.

ii. Qualitative Data Analysis:

- **Thematic Analysis**: Perform thematic analysis on interview transcripts and open-ended questionnaire responses to identify common themes and patterns. Use coding techniques to categorize and interpret the qualitative data.
- **Triangulation**: Combine insights from interviews, questionnaires, and case studies to triangulate findings and enhance the validity and reliability of the qualitative analysis.

3.4. Data Analysis :

3.4.1. Data Analysis Technique

Thematic analysis will be employed to analyze the data collected from semi-structured interviews and document analysis. Thematic analysis involves identifying, analyzing, and reporting patterns within the data to develop themes related to the effectiveness of the ML-IDS (Braun & Clarke, 2006).

3.4.2. Tools

Qualitative data analysis software, such as NVivo or ATLAS.ti, will be utilized to facilitate systematic and rigorous analysis of the data. These software tools offer features for organizing, coding, and analyzing qualitative data, thereby enhancing the efficiency and reliability of the analysis process (Bazeley & Jackson, 2013).

3.4.3. Procedure

i. **Data Preparation:** Transcripts of semi-structured interviews and documents obtained from document analysis will be prepared for analysis. Data will be organized and formatted to ensure consistency and ease of analysis.

ii. **Coding:** Initial coding of the data will involve systematically identifying and labeling segments of text related to key concepts and themes. This process will be conducted iteratively, with codes refined and revised as new insights emerge.

iii. **Theme Development:** Coded segments of text will be grouped together based on shared meanings and patterns to develop overarching themes. Themes will be identified through a process of constant comparison and reflection on the data.

iv. **Data Interpretation:** Once themes have been identified, they will be interpreted in relation to the research questions and objectives. The significance of each theme will be explored, and connections between themes will be examined to develop a comprehensive understanding of the data.

3.4.4. Validation

To ensure the trustworthiness and reliability of the findings, multiple strategies will be employed, including member checking, peer debriefing, and triangulation of data sources (Creswell & Creswell, 2017). Member checking involves verifying the accuracy of interpretations with participants, while peer debriefing involves seeking feedback from colleagues to ensure objectivity and rigor. Triangulation of data sources involves comparing and contrasting findings from different data sources to corroborate conclusions and enhance the credibility of the analysis.

4. Delimitations :

While this study aims to provide valuable insights into the development and evaluation of a machine learning-based intrusion detection system (ML-IDS), it is essential to acknowledge certain delimitations that may influence the scope and generalizability of the findings.

4.1 Scope

This study will focus on the development and evaluation of an ML-IDS within the context of a specific organization. As such, the findings may not be directly applicable to all organizations or industries. However, by focusing on a specific context, the study aims to provide detailed insights into the practical application of ML-IDS in real-world settings (Chauhan & Singh, 2020).

4.2 Time Frame

The research will be conducted over a six-month period, which may limit the observation of long-term impacts and effectiveness of the ML-IDS. While efforts will be made to assess the immediate benefits of the ML-IDS, the study may not capture its long-term performance or scalability over time (Mishra et al., 2019).

4.3 Geographical Focus

The study will be conducted within organizations located in a specific region, which may influence the types of cyber threats encountered and the generalizability of the findings. Different regions may face unique cybersecurity challenges, and as such, the findings of this study may not be directly applicable to organizations operating in other geographical locations (Sowmya & Mary, 2023).

4.4 Organizational Context

The effectiveness of the ML-IDS may be influenced by the specific organizational context, including the size, industry, and existing cybersecurity infrastructure of the participating organizations. While efforts will be made to select organizations representing diverse contexts, the findings may be limited to the specific organizational settings included in the study (Jemili, Rahma, & Quajdi, 2023).

4.5 Technical Limitations

The development and evaluation of the ML-IDS may be constrained by technical limitations, such as resource constraints, technological dependencies, and compatibility issues with existing systems. While every effort will be made to address these limitations, they may impact the implementation and performance of the ML-IDS in practice (Chauhan & Singh, 2020).

5. Ethical Considerations :

Ensuring ethical integrity is crucial in conducting research, especially in sensitive areas like cybersecurity. In this study, ethical considerations will be carefully addressed to uphold the rights and well-being of all participants involved.

5.1 Informed Consent

We value the autonomy of our participants and recognize the importance of informed consent (Chauhan & Singh, 2020). Participants will receive clear and comprehensive information about the research aims, procedures, potential risks, and benefits. Their participation will be entirely voluntary, and they will have the freedom to withdraw from the study at any time without consequences. Signed consent forms will document their informed consent, ensuring transparency and accountability throughout the research process (Mishra et al., 2019).

5.2 Confidentiality and Anonymity

Protecting the privacy and confidentiality of participants' data is of utmost importance (Sowmya & Mary, 2023). All data collected during the study will be anonymized to safeguard participants' identities. Additionally, data will be stored securely on restricted-access servers to prevent unauthorized disclosure or access. These measures ensure that participants' information remains confidential, and their privacy is respected throughout the research process (Chauhan & Singh, 2020).

5.3 Minimization of Harm

We are committed to minimizing any potential harm or discomfort to our participants (Mishra et al., 2019). Care will be taken to ensure that participants are not subjected to undue stress or emotional distress during the study. Additionally, sensitive topics related to cybersecurity will be approached with caution to minimize the risk of causing harm. Participants will have the option to skip or avoid questions that they find uncomfortable, ensuring their well-being and comfort throughout the research process (Sowmya & Mary, 2023).

5.4 Transparency and Integrity

Transparency and integrity are fundamental principles guiding our research conduct (Jemili, Rahma, & Quajdi, 2023). We are committed to conducting the research with honesty, accuracy, and openness. Data collected during the study will be accurately documented and preserved to ensure transparency and reproducibility. Any conflicts of interest will be disclosed, and the research findings will be reported honestly and transparently, without misrepresentation or manipulation (Chauhan & Singh, 2020).

5.5 Compliance with Regulations

We will adhere to all relevant ethical guidelines, regulations, and institutional policies governing research conduct (Sowmya & Mary, 2023). Necessary ethical approvals will be obtained from the appropriate ethics review board before the commencement of the research. Additionally, the research will comply with data protection and privacy regulations to safeguard participants' personal information. These measures ensure that the research is conducted ethically and in compliance with established standards and regulations (Jemili, Rahma, & Quajdi, 2023).

6. Conclusions :

In conclusion, this research aims to address the need for innovative solutions to combat modern cyber threats by developing and evaluating a machine learning-based intrusion detection system (ML-IDS). Through a comprehensive research design, including a case study approach, semi-structured interviews, and document analysis, this study will explore the effectiveness and practical value of ML-IDS in protecting against cyber threats. Ethical considerations will be carefully adhered to throughout the research process to ensure the well-being and rights of all participants. This study seeks to contribute to the advancement of cybersecurity by providing organizations with a robust and adaptive defense mechanism against a wide range of cyber threats.

7. References :

i. Ahmad, I., Basheri, M., Iqbal, M. J., & Rahim, A. (2018). Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection. *IEEE Access*, 6, 33789-33795.

ii. Chauhan, A., & Singh, A. (2020). Machine Learning-based Intrusion Detection Systems: A Comprehensive Survey. *Future Internet*, 12(7), 117.

iii. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE Publications.

iv. Gibbs, G. R. (2018). *Ethical Practice in Social Research.* SAGE Publications.

v. Mishra, R., Jain, N., Singh, S., & Mohapatra, D. P. (2019). Machine Learning Techniques for Intrusion Detection Systems: A Review. *Artificial Intelligence Review*, 52(3), 1983-2009.

vi. Resnik, D. B. (2015). *What is Ethics in Research & Why is it Important?* National Institute of Environmental Health Sciences.

vii. Sowmya, K. R., & Mary, J. A. (2023). A Review on Intrusion Detection Systems Using Machine Learning Techniques. *International Journal of Scientific & Technology Research*, 2(9), 204-208.

viii. Jemili, I., Rahma, B., & Quajdi, H. (2023). *Intrusion Detection Systems: State of the Art and New Perspectives*. International Journal of Computer Applications, 180(48), 29-34.

ix. Cybersecurity and Infrastructure Security Agency. (2021). *Critical Infrastructure Security*. Retrieved from https://www.cisa.gov/.

x. European Union Agency for Cybersecurity. (2020). *Threat Landscape Report 2020*. Retrieved from https://www.enisa.europa.eu/publications/threat-landscape-report-2020

xi. Mansfield-Devine, S. (2019). *Implementing Cybersecurity: A Guide to the National Institute of Standards and Technology Risk Management Framework*. Apress.

xii. National Institute of Standards and Technology. (2018). *Computer Security Incident Handling Guide*. Retrieved from https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf