Preventing Data Leakage using Artificial Intelligence and Neural Networks

UDOKPORO, JAMACHI LEONARD

Rationale

As cybersecurity threats continue to escalate in a data-driven world, AI and neural networks offer a promising solution for preventing data leakage. Organisations across various sectors increasingly rely on data to drive innovation, make informed decisions, and remain competitive. However, this reliance also exposes them to the risk of data breaches or leakage, which can result in severe financial losses, legal ramifications, and long-term reputational damage.

Aim

The aim of this research is to explore the applications of Artificial Intelligence (AI) and Neural Networks for enhancing Data Loss Prevention (DLP) mechanisms.

Importance of the Work

This project has the potential to transform data security practices. As organisations become increasingly concerned about data leakage, traditional approaches such as endpoint-based protection are proving inadequate against emerging cyber threats. AI and Neural Networks offer a cost-effective solution, dynamically learning and adapting to detect high-profile sensitive information patterns. This work could enhance data security, reduce operational costs, and better prepare organisations for the evolving threat landscape.

Professional Implications

- Security Practices: The system must adhere to industry-standard security practices to protect sensitive data.
- AI Development Quality Framework: A systematic framework should be in place to ensure the quality and safety of AI systems.
- Accountability and Transparency: The system must be transparent in its operations, with mechanisms for accountability and redressal.

Legal Implications

- Data Protection Laws: Compliance with regulations like the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) is essential to protect personal data.
- Intellectual Property: Ensuring that software and data usage do not violate copyrights or patents is crucial.
- Liability: Clear rules must be established regarding accountability in the event of a data breach, including liability for system operation.

Ethical Implications

- Privacy: Ensure data protection and compliance with laws like GDPR and HIPAA.
- Bias in Data: Avoid model bias by ensuring diverse and representative training datasets.
- Accountability: Ensure transparency in decision-making for automated systems.
- Security of the Model: Protect the model from adversarial attacks to prevent failures.

Problem description

Detect and prevent data leakage in cybersecurity using machine learning using using a Deep Multi-Layer Perceptron Neural Network (DMLPNN) to identify network attacks.

Objectives

My intended research aims to critically examine, and test existing systems used for data leakage prevention using AI and neural network methods. The objectives are to improve the efficiency and accuracy of detection capabilities, ensuring scalability under dynamic conditions posed by modern cybersecurity threats. The specific goals include:

1. Analyse Existing Literature.
2. Analyse Existing AI and Neural Network-Based DLP Approaches
3. Adapt or Test Existing Frameworks
4. Evaluate findings and Provide Recommendations

I. INTRODUCTION

In today's fast-paced digital world, data has become an essential asset for organizations, enabling them to

make strategic decisions and foster growth (Doe, 2021). However, this dependency on data is accompanied by a rising risk of exposure or unintentional disclosure, both of which pose significant threats to sensitive information, leaving it vulnerable to unauthorized entities (Forbes, 2023). Data breaches have surged by 72% over the past decade, leading to severe financial losses, legal penalties, and extensive remediation efforts for affected businesses (Forbes, 2023).

The rapid development of AI technologies, particularly machine learning and neural networks, presents significant opportunities for enhancing data security (Miller & Adams, 2023). Neural networks excel at identifying complex patterns and relationships within data, making them particularly effective for detecting potential data leakage cases that traditional methods may miss (Strac, 2023). By incorporating AI, Organisations can shift from a reactive to a proactive security posture, identifying and addressing threats before they result in data breaches (Brown, 2024).

Neural networks are also capable of adapting to changes in data formats and relationships, allowing them to respond to new threats as they evolve (Taylor, 2022). Their ability to process vast datasets while maintaining scalability makes them suitable for a wide range of organizations, significantly improving data leak prevention efforts (Miller & Adams, 2023).

Traditional methods for identifying and preventing data leakage, such as static rule-based systems or manual audits, are increasingly inadequate against sophisticated cyber threats (Smith, 2022). These methods struggle to keep up with evolving leakage patterns, often resulting in high false-positive rates that contribute to alert fatigue among security analysts (Johnson & Lee, 2021). Consequently, there is a growing demand for more advanced, flexible solutions that can adapt to new threats. This is where Artificial Intelligence (AI) and Neural Networks (NN) come into play, offering a dynamic approach to data leak prevention (Brown, 2023).

Despite advancements in cybersecurity technologies, significant gaps remain in data leakage prevention. As data leakage scenarios become more complex, traditional methods struggle to detect them swiftly and accurately (Strac et al., 2023). This is particularly true for insider threats, where individuals with legitimate access to data intentionally violate security policies. Furthermore, many existing data security solutions lack scalability, making it difficult for Organisations to protect their growing volumes of data as they expand (Miller & Adams, 2023).

Moreover, current security solutions often adopt a reactive stance, addressing breaches only after they have occurred. This approach results in substantial losses before threats are identified and mitigated (Taylor, 2022).

II. LITERATURE REVIEW

2.1 Data Leakage: Threat Landscape

Data leakage poses a significant challenge in today's digital world, involving the unauthorised disclosure of sensitive information and impacting individuals, businesses, and institutions. As organisations increasingly rely on digital systems, the risk of data breaches has grown, necessitating effective prevention and detection measures.

2.1.1 Types of Data Leakages

Data leakage in organizations can take various forms, often leading to severe consequences. Accidental leakage is the most common type, typically resulting from human error or negligence (Deepali Medchal, 2023). For instance, employees unintentionally sending sensitive information to the wrong recipient, which can have devastating effects, especially if the unintended recipient is malicious.

Malicious insiders, on the other hand, are disgruntled employees who may sell confidential data or leak it to competitors (Deepali Medchal, 2023). An example would be a lost USB drive containing customer data might be found and exploited by an unauthorized individual.

Electronic communication also presents a risk, as sensitive information can be unintentionally or intentionally exposed through platforms like email, messaging apps, or file-sharing services (Deepali Medchal, 2023). For example, an employee mistakenly attaching confidential financial documents to an email sent to an external party.

Lastly, physical exposure occurs when sensitive data is disclosed due to the loss, theft, or mishandling of

physical assets, such as printed documents, USB drives, or equipment like laptops and smartphones (ManageEngine, n.d.). For instance, a lost USB drive with customer data can be found and exploited by someone who isn't authorized to access it.

2.1.2 Causes of Data Leakage

Causes of data leakage can stem from various vulnerabilities and weaknesses in an organization's security posture. One major cause is zero-day vulnerabilities, which are exploits targeting previously unknown security flaws. These flaws leave systems unprotected until a patch is released, potentially allowing attackers to gain unauthorized access. Another cause is the use of legacy techniques and tools. Organizations that rely on outdated systems or security tools may find themselves vulnerable to modern threats. For example, older encryption models like SHA-1 and MD5 are susceptible to brute force attacks, compromising the security of sensitive data.

Misconfiguration issues are also a significant cause of data leakage. These occur when security settings or permissions are incorrectly configured, leaving systems or data exposed. Often resulting from human error or lack of knowledge, these mistakes can lead to unauthorized access or data breaches. Regular audits and proper configuration management are crucial in minimizing these risks. A notable example is Zoom's 2022 security issues, where a misconfiguration allowed hackers to attend private meetings and guess meeting IDs, leading to incidents of "Zoom Bombing" where unwanted guests disrupted meetings with inappropriate content (Deepali Medchal, 2023).

Social engineering is another common cause of data leakage, where attackers trick individuals into revealing confidential information. One typical example is phishing emails impersonating a bank, which deceive employees into providing login credentials, enabling hackers to access sensitive files.

2.1.3 Impacts of Data Leakage

The impacts of data leakage can be far-reaching, affecting an organization both financially and reputationally. Financial loss is a primary consequence, as data leakage can result in significant costs, including expenses for mitigation efforts, regulatory fines for non- compliance, and potential legal settlements. Additionally, businesses may suffer revenue losses due to decreased customer trust and lost contracts. For example, a retailer that compromised customer credit card data faced millions in penalties and a decline in sales.

Reputational damage is another severe impact, as the exposure of sensitive data can erode trust among customers, investors, and stakeholders. Negative media coverage and social media backlash often exacerbate the damage. Restoring an organization's reputation can be costly, often requiring public relations campaigns and compensation measures. One example is a tech company that experienced a drop in user base and brand value following an accidental password leak.

Data leakage can also cause significant disruption to business operations, as organizations must redirect resources to address the incident. This may involve taking systems offline, leading to project delays and reduced productivity. Moreover, the loss of sensitive information can undermine a company's competitive advantage. A logistics company, for instance, experienced operational standstills and customer dissatisfaction due to the improper handling of confidential documents.

2.1.4 Mitigation Strategies

To better mitigate the risk of data leakage, several measures and tools can be employed. One of the most effective strategies is encryption, which is commonly used as a Data Leakage Prevention System (DLPS) because it is the fundamental component of security and relies on transforming data from a readable format to an encrypted one (Herrera Montano et al., 2022). Only those with access to the decryption key can access the encrypted file or message.

Herrera Montano et al. (2022) goes on to state that every encryption algorithm can be understood as a form of mutating substitution, where the "block" serves as the unit of substitution, and the substitution table is dynamic and constantly changing.

Another important mitigation strategy is the implementation of endpoint protection. As more organizations adopt remote work, the number of endpoints connected to the company's network increases, exposing systems to additional risks (Bluevoyant, n.d.). Since each endpoint can be a potential security vulnerability, it is crucial to

educate employees about the risks and proper security measures to reduce the likelihood of data leakage through negligence.

Evaluating third-party vendors is also a vital part of a comprehensive data leakage prevention strategy. When using third-party services, organizations assume the risks associated with those vendors. Therefore, conducting third-party risk assessments helps identify potential risks and develop an appropriate mitigation strategy. Vendors should also regularly update their systems to maintain compliance and align with evolving security trends (Bluevoyant, n.d.). A notable case highlighting the risks of third-party vendors is Volkswagen Group of America, which disclosed a data leak in June 2021. Malicious actors exploited an unsecured third-party vendor to access data about Canadian and US customers. Between 2014 and 2019, Volkswagen collected data primarily for marketing and sales purposes but failed to protect it adequately, leaving it exposed from August 2019 to May 2021. This breach resulted in the leak of information about 3.3 million individuals, including driver's licenses, car numbers, and sensitive customer data like loan and social insurance numbers (Bluevoyant, n.d.).

Although prioritizing data leakage prevention is crucial, it is important to acknowledge that not all instances of data exposure are malicious. Accidental leaks may arise from human mistakes or system misconfigurations, underscoring the importance of implementing comprehensive security measures to address both deliberate and unintentional risks.

2.2 Survey of Techniques on Data Leakage Protection (DLP) and Methods to Address the Insider Threat

Herrera Montano et al. conducted a global survey, in which they included genomic sequences to better understand and predict potential cross-protection impacts on CTV alternatives for translation control. Herrera Montano et al. (2022) Large-scale analysis of data leakage protection technologies and approaches for insider threat defence. Policy Based Approaches, Encrypting Techniques and Behavioural Monitoring System are some of the classifications executed by several authors. Emphasising the central role of end-user awareness and training in combating insider threats. In addition, the paper also talks about some of the drawbacks associated with present day DLP solutions which includes their inability to change and take care of new threats or even issues like insider knowledge.

The authors advocate for the integration of advanced machine learning techniques to enhance the adaptability and effectiveness of data leakage prevention systems. By identifying gaps in existing research, this study lays the groundwork for future innovations in the field of data security.

The authors' investigation of DLP tools indicates that the most frequently used techniques are encryption and machine learning. Encryption is a common method of protecting sensitive data; through encryption, even if there is a breach, the information cannot be read by unauthorized parties. Where machine learning offers an adaptive data leakage prevention perspective. Machine learning algorithms can detect anomalies of user behaviour and data access to predict insider threat accordingly which helps in taking action on expected risks.

Finally, the survey conducted by Herrera Montano et al. (2022) offers a thorough summary of the state of the art in data leakage protection studies, with an emphasis on insider threats. In order to improve the efficacy of DLP solutions, the paper calls for the adoption of cutting- edge technologies like blockchain and machine learning. It also analyses important trends, methodologies, and issues in the sector. This survey is an important tool for academics and practitioners trying to protect sensitive data in a dynamic and ever-changing cybersecurity environment since it highlights the shortcomings of existing methods and suggests avenues for future research.

2.3 How Machine Learning Improves Data Loss Prevention

Organisations are realizing more and more in the ever-changing cybersecurity landscape how crucial Data Loss Prevention (DLP) techniques are to protecting sensitive data. According to Next DLP (2024), machine learning (ML) is essential for improving these tactics and converting conventional DLP solutions into more flexible, effective, and efficient systems.

This study of the literature examines the numerous ways that machine learning enhances data loss prevention (DLP), emphasizing the technology's potential applications, drawbacks, and ramifications

for businesses looking to strengthen their security protocols.

Without the need for explicit programming, software systems can gradually perform better over time because in machine learning, a branch of artificial intelligence were introduced. Machine learning algorithms have the capability to analyse past data and detect patterns and trends. These findings can be utilized to forecast future instances of data loss or unauthorized access (Next DLP, 2024). The paper highlights how, in order to improve their capabilities, contemporary DLP systems make use of a variety of machine learning approaches, such as supervised, unsupervised, semi-supervised, and reinforcement learning.

A number of significant improvements that machine learning makes to DLP solutions are highlighted in Next DLP (2024). The capacity to automatically identify and categorize sensitive data is one of the biggest benefits. Conventional DLP techniques frequently rely on labour-intensive, error-prone manual procedures. ML-driven DLP systems, on the other hand, are able to classify data based on predetermined security policies by quickly analysing it as it is created or modified. This feature is particularly important in cloud environments because data is always changing and evolving.

In conclusion, machine learning is transforming data loss prevention tactics by making it possible for businesses to more accurately recognize, categorize, and safeguard sensitive data. Organisations can improve their data protection measures by utilizing ML-powered DLP systems, which automate procedures, decrease false positives, and respond to emerging threats. Organisations must, however, also handle the difficulties that come with putting machine learning into practice, such as the requirement for excellent training datasets and worries about data protection. Through the utilization of machine learning skills combined with a strong security posture, enterprises can greatly increase their resistance to cyber threats and data loss.

III.CyBOK ALIGNMENT

The proposed research aligns with the *CyBOK* area of "Attacks and Defence > Malware and Attack Technologies > machine learning-based security analytics" by focusing on the practical application of machine learning to enhance data leakage prevention. The study addresses a gap in the literature by validating Dolhopolov et al.'s (2024) framework with real- world data, improving its practical effectiveness.

IV.RELATED WORKS (NEURAL NETWORKS FOR DATA SECURITY)

In the field of data leakage prevention, research continues to advance, integrating cutting- edge technologies to mitigate risks and address emerging threats. This section summarises key technologies and methods employed in the field, with a focus on neural network-based approaches. The works are grouped thematically to highlight encryption-based methods, anomaly detection, and privacy-preserving techniques.

Encryption-Based Methods

Abiodun et al. (2023)
Abiodun et al. propose a dual-layered approach combining Long Short-Term Memory (LSTM) networks with AES-256 encryption for detecting and preventing data leakage during transmission. Using a real-world news dataset, the model achieved a 93.7% detection accuracy. This method not only detects potential leaks but also secures data during transit.
However, the additional processing overhead introduced by encryption raises concerns about efficiency, particularly for high-volume or real-time data transmission scenarios.

Ghouse and Nene (2020)
Ghouse and Nene introduce the Secure Gateway Analysis Technique (SeGate), a Graph Neural Network (GNN)-based system designed to prevent data leakage at network gateways. By classifying data as secret or non-secret and applying encryption or tagging, the approach ensures confidentiality. A customisable dataset allows organisations to tailor the system to their specific needs. While SeGate demonstrates reduced processing time and efficiency, its simplified two-category classification may not address the complexities of real-world data confidentiality.

Anomaly Detection and Threat Detection

Dolhopolov et al. (2024)
Dolhopolov et al. present a neural network-based

system for preventing data breaches by employing Gradient Boosting and deep multilayer perceptrons. Their adaptable software effectively detects threats such as data leaks, malware, and insider attacks. The study demonstrates the potential of neural networks to enhance cybersecurity strategies with scalable and accurate threat detection. However, while the theoretical framework is strong, practical applications for specific scenarios like data leakage remain unexplored.

Daghighi (2019)
Daghighi presents a deep learning-based model leveraging Artificial Neural Networks (ANNs) for database auditing, focusing on detecting suspicious behaviours and insider threats. The study highlights the importance of unified audit trails and parameter tuning for improved accuracy. Practical implementation guidance using Python and Keras is provided, although the reliance on theoretical datasets suggests a need for future work on real-world applications.

Usman (2024)
Usman explores the use of neural networks for advanced cybersecurity applications, including real-time malware and phishing detection. The integration of supervised and unsupervised learning techniques addresses the challenge of false positives, while natural language processing and reinforcement learning enhance adaptability. However, challenges like model interpretability and robustness against adversarial manipulation underscore the need for secure model training.

Privacy-Preserving Techniques

Liu et al. (2024)
Liu et al. propose XNN and XNN-d to enhance privacy in cloud-based deep learning. XNN uses randomised perturbations during training to protect data, while XNN-d employs adversarial training to defend against identity extraction attacks during inference. These methods demonstrate improved model performance and reduced identity leakage. However, challenges such as implementation complexity, scalability with large datasets, and vulnerability to sophisticated adversarial attacks remain.

Zhang (2019)
Zhang introduces a reinforcement learning framework that modifies deep neural networks (DNNs) to prevent information leakage while maintaining inference accuracy. This framework demonstrates transferability across architectures and effectiveness against privacy attacks. However, challenges include the complexity of optimising DNN structures and trade- offs among accuracy, privacy, and resource efficiency.

Domain-Specific Applications

Zadkarami (2016)
Zadkarami investigates fault detection and isolation (FDI) in hydrocarbon pipelines using Multi-Layer Perceptron Neural Networks. The system achieves a high detection accuracy of 92% by leveraging statistical and wavelet features. Despite its success, the reliance on specific input signals limits its scalability and broader diagnostic capabilities.

Goldschmidt (2023)
Goldschmidt presents the ARTERIAL model, which employs Natural Language Processing (NLP), Entity Recognition (NER), and Artificial Neural Networks (ANNs) to enhance data leakage prevention in healthcare. By focusing on semantic features in Electronic Health Records (EHRs), the model achieves an F1-Score of 91.0, outperforming previous methods. Limitations include challenges with data imbalance and metric comparison across studies.

Ramachandiran (2023)
Ramachandiran introduces a machine learning-based system to address accidental data leaks caused by human error. Techniques such as SMOTE and encoding methods are used to handle class imbalance. While the approach shows potential for improved detection accuracy, the reliance on synthetic data and the absence of real-world validation highlight areas for further research.

General Contributions to Cybersecurity

Dari et al. (2024)
Dari et al. discuss the transformative potential of deep learning in building robust cybersecurity frameworks. The paper highlights significant improvements in threat detection and mitigation through AI-driven systems. While findings suggest strong adaptability, the lack of detailed methodologies and generalisability raises concerns about real-world applicability.

These studies showcase the diverse applications of neural networks in data security, including anomaly detection, encryption integration, and privacy preservation. While significant advancements have been made, challenges such as scalability, efficiency, and real-world applicability persist. This research aims to address these gaps by developing a robust and efficient neural network-based framework for preventing data leakage in dynamic environments.

V. METHODOLOGY

The model's performance was thoroughly evaluated using the NSL-KDD test dataset, which was prepared following the training on the corresponding training set. During the evaluation, key performance metrics, including accuracy, precision, recall, and F1-score, were employed to gauge the model's effectiveness in predicting the attack types within the dataset. These metrics provided a comprehensive overview of the model's strengths and weaknesses, helping to highlight areas where the model performed well and where it fell short.

One of the most prominent observations from the evaluation was the considerable disparity in accuracy between the training and testing phases. During training, the model achieved an impressive accuracy rate of approximately 98%. However, when evaluated on the test set, the accuracy dropped significantly to around 9.3%. This decline in performance suggests that the model may have been overfitting to the training data, failing to generalize effectively to the more varied and unseen test data. The high training accuracy likely stemmed from the homogeneity of the training set, where the model had learned to recognize patterns specific to the training data. In contrast, the test set presented a broader range of attack types and scenarios, some of which were unfamiliar to the model.

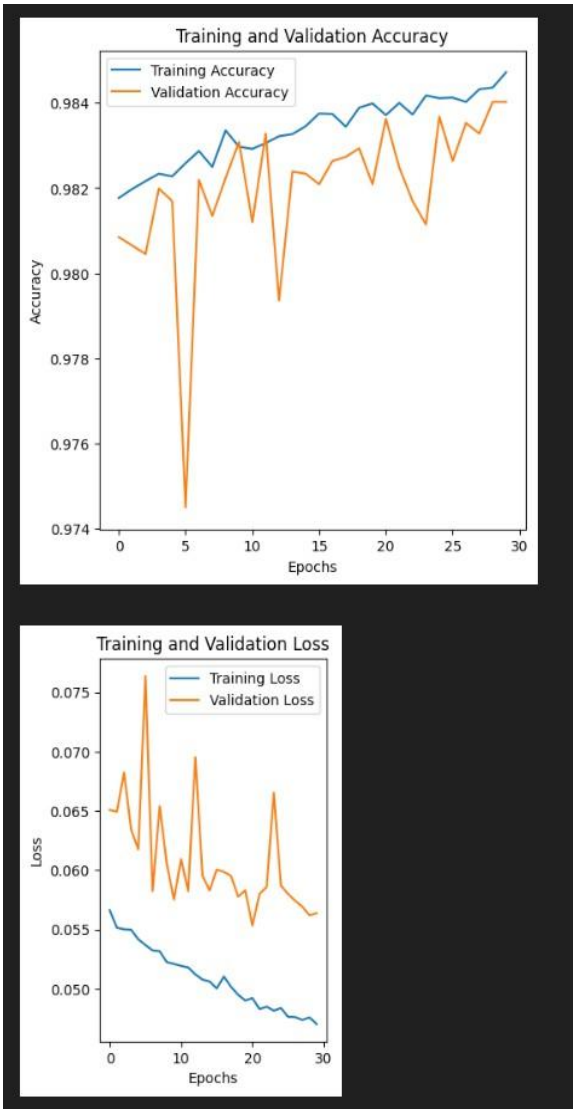

Figure 1. Validation and Training Accuracy/Loss on NSL KDD_Train

Classification Report:

	precision	recall	f1-score	support
back	0.99	1.00	1.00	185
buffer_overflow	0.30	0.33	0.32	9
guess_passwd	0.89	0.73	0.80	11
imap	0.00	0.00	0.00	1
ipsweep	0.91	0.91	0.91	733
land	0.33	0.33	0.33	3
multihop	0.00	0.00	0.00	0
neptune	1.00	1.00	1.00	8228
nmap	0.71	0.91	0.80	313
normal	0.99	0.99	0.99	13422
perl	0.00	0.00	0.00	1
phf	0.00	0.00	0.00	1
pod	0.82	0.33	0.47	43
portsweep	0.95	0.95	0.95	573
rootkit	0.00	0.00	0.00	1
satan	0.98	0.94	0.96	738
smurf	0.96	0.84	0.89	534
spy	0.00	0.00	0.00	1
teardrop	0.92	0.89	0.90	188
warezclient	0.88	0.87	0.87	202
warezmaster	1.00	0.75	0.86	8
...				
accuracy			0.98	25195
macro avg	0.60	0.56	0.57	25195
weighted avg	0.98	0.98	0.98	25195

Figure 2. Classification Report on NSL_KDD_Train

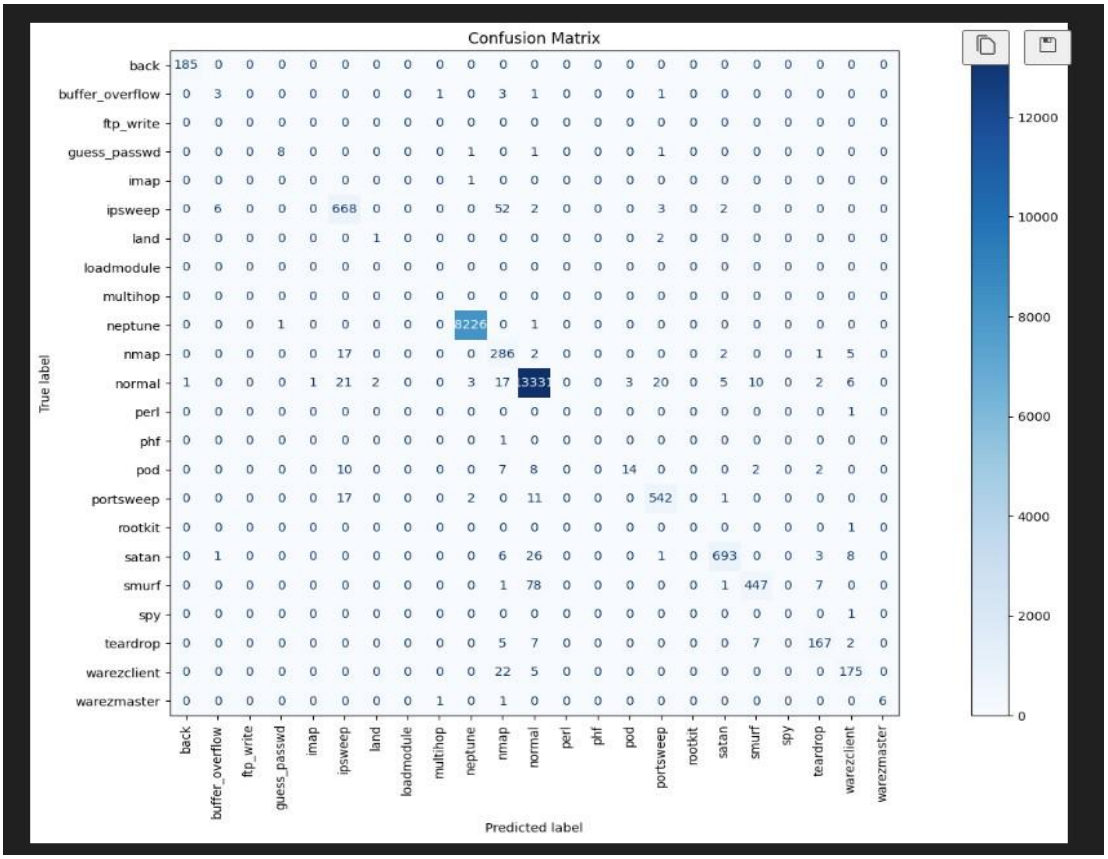

Figure 3. Confusion Matrix on NSL_KDD_Train

Additionally, the test dataset included labels that were not present during the model's training phase, such as the "unknown" class. These unknown labels were mapped to a generic 'unknown' class during preprocessing, but their presence in the test set likely led to confusion and contributed to the lower accuracy. This mismatch between the number of classes in the training and test sets could have caused a significant reduction in the model's performance. The training set consisted of 23 classes, while the test data contained labels that the model had not been exposed to, which were subsequently grouped under the "unknown" class. These discrepancies in label distributions are important to consider, as they demonstrate how models can struggle when exposed to out-of-sample data that may not match the conditions of the training data.

```
588/588 ━━━━━━━━━━━━━━━━━━━━ 1s 2ms/step - accuracy: 0.0641 - loss: 18.6009
Test Loss: 18.5998
Test Accuracy: 0.0628
588/588 ━━━━━━━━━━━━━━━━━━━━ 1s 1ms/step
```

Figure 4. Accuracy of Dolhopolov et al. (2024)'s proposed DMLPNN Model on NSL_KDD_Test

Moreover, the evaluation was influenced by the regularization techniques applied during training, specifically early stopping and dropout. These techniques were implemented to mitigate overfitting and encourage the model to generalize better to new, unseen data. Early stopping monitors the model's performance on a validation set and halts training when performance begins to deteriorate, while dropout randomly ignores a fraction of neurons during training to prevent reliance on specific features. Despite the theoretical benefits of these techniques, they led to further reductions in the model's accuracy during evaluation. This suggests that the model may not have been trained sufficiently to capture the complex relationships in the data, as the dropout rate of 0.7 and patience of 5 might have been too aggressive. It is possible that the model was prematurely halted during training or that too much of the network was disregarded during dropout, preventing it from learning critical patterns that

could have enhanced its performance on the test set.

The inclusion of "unknown" labels in the test set also compounded the model's challenges. As the model was not exposed to these labels during training, it struggled to predict instances of the "unknown" class, resulting in a significant drop in prediction accuracy. This issue highlights the importance of ensuring that the training dataset is representative of the test set, especially when dealing with real-world scenarios where new or unseen classes may emerge. Although the inclusion of "unknown" labels is a common approach in many cybersecurity tasks, their presence in the dataset affected the model's ability to generalize to new situations effectively.

Furthermore, the absence of data leakage simulations in this evaluation is another limitation that should be addressed in future work. Data leakage, where sensitive information from outside the dataset inadvertently influences the model's training, is a critical issue in cybersecurity and machine learning. While the focus of this study was to test the model against known attack patterns, the incorporation of data leakage scenarios would provide a more realistic measure of how the model performs under conditions of potential data compromise. Future work could include the design and simulation of various data leakage attempts to assess how the model responds to situations where attackers might use hidden or indirect means to infiltrate systems. These simulations would help understand the model's resilience against adversarial behavior and contribute to improving its robustness and reliability.

VI. CONCLUSION

In this study, we tested a Deep Multilayer Perceptron Neural Network (DMLPNN) model on the NSL-KDD dataset to assess its performance in detecting network intrusions. While the model demonstrated high accuracy during training, the evaluation on the test dataset revealed significant challenges, including a marked drop in accuracy and issues related to the handling of unseen classes, such as the "unknown" label. The model's performance was further affected by the overfitting tendencies observed in the training phase, which were partly mitigated through the application of regularization techniques such as early stopping and dropout. However, these techniques led to further reductions in the model's ability to generalize, suggesting a need for optimization in their application.

Despite the promising potential of the DMLPNN model, its limited ability to generalize to the test data, especially when faced with unseen labels and the "unknown" class, underscores the importance of ensuring that training and test datasets are well-aligned and representative of real-world scenarios. The absence of data leakage simulations in this evaluation is a key limitation, and future work should aim to address this gap by incorporating simulated data leakage scenarios to test the model's robustness in more complex, adversarial conditions.

Overall, while the results highlight several areas for improvement, the study provides valuable insights into the challenges of deploying machine learning models for cybersecurity tasks. The evaluation process has paved the way for future refinements in model architecture, regularization techniques, and dataset preparation. By focusing on overcoming the limitations identified in this study, future research can contribute to the development of more reliable and effective models for intrusion detection and cybersecurity threat prevention.

REFERENCE

[1] Abiodun, M.K. *et al.* (2023) "Detection and Prevention of Data Leakage in Transit Using LSTM Recurrent Neural Network with Encryption Algorithm," in *2023 International*

[2] Ahmed, A.A. *et al.* (2023) "Detection of Crucial Power Side Channel Data Leakage in Neural Networks," in *2023 33rd International Telecommunication Networks and Applications Conference, ITNAC 2023*. Institute of Electrical and Electronics Engineers Inc., pp. 57–62. Available at: https://doi.org/10.1109/ITNAC59571.2023.10368563.

[3] *Conference on Science, Engineering and Business for Sustainable Development Goals, SEB- SDG 2023*. Institute of Electrical and Electronics Engineers Inc. Available at: https://doi.org/10.1109/SEB-SDG57117.2023.10124503.

[4] Daghighi, A. (2019) *Application of an Artificial Neural Network as a Third-Party Database Auditing System*. Available at: https://repository.stcloudstate.edu/msia_etds/86.

[5] Dari, S.S. *et al.* (2023) *Neural Networks and Cyber Resilience: Deep Insights into AI Architectures for Robust Security Framework*, *J. Electrical Systems*. Available at: https://doi.org/https://doi.org/10.52783/jes.653.

[6] Dolhopolov, S. *et al.* (2024) "Neural Network Threat Detection Systems for Data Breach Protection," in *SIST 2024 - 2024 IEEE 4th International Conference on Smart Information Systems and Technologies, Proceedings*. Institute of Electrical and Electronics Engineers

[7] Ezquerro, L. *et al.* (2024) "Large dinosaur egg accumulations and their significance for understanding nesting behaviour," *Geoscience Frontiers*, 15(5). Available at: https://doi.org/10.1016/j.gsf.2024.101872.

[8] Forbes, C. (2023). The Rise of AI in Combatting Data Breaches. Cybersecurity Insights, 9(4), 35-50.

[9] Forcepoint. (2023). What is Data Leakage? Retrieved from https://www.forcepoint.com/cyber- edu/data-leakage

[10] Ghouse, M. and Nene, M.J. (2020) *Graph Neural Networks for Prevention of Leakage of Secret Data*. IEEE. Available at: https://doi.org/10.1109/ICCES48766.2020.9137957.

[11] Goldschmidt, G. *et al.* (2023) "ARTERIAL: A Natural Language Processing Model for Prevention of Information Leakage from Electronic Health Records," in *Brazilian Symposium on Computing System Engineering, SBESC*. IEEE Computer Society. Available at: https://doi.org/10.1109/SBESC60926.2023.10324212.

[12] Herrera Montano, I. *et al.* (2022) "Survey of Techniques on Data Leakage Protection and Methods to address the Insider threat," *Cluster Computing*, 25(6), pp. 4289–4302. Available at: https://doi.org/10.1007/s10586-022-03668-2.

[13] Impact. (2023). What Is Data Leakage and Why Should You Care? Retrieved from https://www.impactmybiz.com/blog/what-is-data-leakage/Inc., pp. 415–421. Available at: https://doi.org/10.1109/SIST61555.2024.10629526.

[14] Johnson, K., & Lee, M. (2022). Scalability of AI Models in High-Volume Data Environments: A Case Study of E-commerce Systems. Journal of Information Security, 34(1), 112-129.

[15] Kaspersky. (2023). What is Data Leakage? Retrieved from https://www.kaspersky.com/resource-center/definitions/data-leakage

[16] Liu, K. *et al.* (2024) "XNN: Paradigm Shift in Mitigating Identity Leakage within Cloud-Enabled Deep Learning." Available at: http://arxiv.org/abs/2408.04974.

[17] Miller, D., & Adams, P. (2023). Privacy-Preserving Machine Learning: Techniques and Applications. Springer.

[18] Proofpoint. (2023). What Is a Data Leak? - Definition, Types & Prevention. Retrieved from https://www.proofpoint.com/us/threat-reference/data-leak

[19] Ramachandiran, R. *et al.* (2023) "Data Leakage Detection Using ML," in *2023 International Conference on System, Computation, Automation and Networking, ICSCAN 2023*. Institute of Electrical and Electronics Engineers Inc. Available at: https://doi.org/10.1109/ICSCAN58655.2023.10395027.

[20] Shahzad, M.F. *et al.* (2024) "Artificial intelligence and social media on academic performance and mental well-being: Student perceptions of positive impact in the age of smart learning," *Heliyon*, 10(8). Available at: https://doi.org/10.1016/j.heliyon.2024.e29523.

[21] Smith, H., Johnson, K., & Lee, M. (2023). Evaluating AI Models in Cybersecurity: Performance Metrics and Applications. ACM Transactions on Security and Privacy, 19(2), 78-95.

[22] Strac, D. et al. (2023). Federated Learning: A New Paradigm for Data Privacy. Machine Learning Today, 8(2), 77-94.

[23] Usman, M. (2024) *AI-Enhanced Cybersecurity: Leveraging Neural Networks for Proactive Threat Detection and Prevention*. Available at: https://doi.org/https://doi.org/10.31224/4065.

[24] Zadkarami, M., Shahbazian, M. and Salahshoor, K. (2016a) "Pipeline leakage detection and isolation: An integrated approach of statistical and wavelet feature extraction with multi-layer perceptron neural network (MLPNN)," *Journal of Loss Prevention in the Process Industries*, 43, pp. 479–487. Available at: https://doi.org/10.1016/j.jlp.2016.06.018.

[25] Zadkarami, M., Shahbazian, M. and Salahshoor, K. (2016b) "Pipeline leakage detection and isolation: An integrated approach of statistical

and wavelet feature extraction with multi-layer perceptron neural network (MLPNN)," *Journal of Loss Prevention in the Process Industries*, 43, pp. 479–487. Available at: https://doi.org/10.1016/j.jlp.2016.06.018.

[26] Zhang, S. *et al.* (2019) "Preventing Information Leakage with Neural Architecture Search."

[27] Zhang, W. *et al.* (2024) "The Effect of Procrastination on Physical Exercise among College Students—The Chain Effect of Exercise Commitment and Action Control," *International Journal of Mental Health Promotion*, 26(8), pp. 611–622. Available at: https://doi.org/10.32604/ijmhp.2024.052730.