International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(RESEARCH ARTICLE)

Check for updates

Analyzing detection algorithms for cybersecurity in financial institutions

Balaji Dhashanamoorthi *

Individual Researcher, Chennai, India.

International Journal of Science and Research Archive, 2024, 11(02), 558–568

Publication history: Received on 08 February 2024; revised on 16 March 2024; accepted on 19 March 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.11.2.0478

Abstract

Frauds in financial services are an ever-increasing phenomenon, and cybercrime generates multimillion revenues. Even a small improvement in fraud detection rates would lead to significant savings. Traditional rule-based systems have limitations in blocking potentially fraudulent transactions. This chapter explores how machine learning, specifically supervised and unsupervised learning, can address these limitations more effectively.

We present a novel AI-based fraud detection system that combines supervised and unsupervised models. In the batch layer, transaction data undergoes pre-processing and model training, while the stream layer handles real-time fraud detection based on new input transaction data. The architecture automates fraud detection processes, making it a valuable tool for supporting fraud analysts.

This research aims to enhance cybersecurity in financial institutes by leveraging the power of AI and machine learning. The integration of supervised and unsupervised models provides a robust defense against cyber faults, ensuring the safety of financial transactions.

Keywords: Artificial intelligence; Fraud detection; Real-time analysis; Machine learning; Automation; Supervised learning; Unsupervised learning

1. Introduction

In today's interconnected digital landscape, financial institutes face relentless cyber threats that jeopardize data security, financial stability, and customer trust. As cybercriminals evolve their tactics, the need for robust defense mechanisms becomes paramount. This journal explores the application of supervised and unsupervised learning techniques to prevent and mitigate cyber faults within financial institutions.

Cyber security is a crucial issue in the modern world, as various cyberspaces are used by criminals to conduct cybercrime and cyber threats. To cope with these challenges, the banking and financial industry has adopted artificial intelligence (AI) as a promising technology that can perform various functions associated with human minds, such as reasoning, learning, interacting, creating, perceiving, and problem-solving. AI can also handle large volumes of structured and unstructured data, extract useful patterns and insights, and control individual human behavior, inference methods, and knowledge representation. However, AI also has some limitations and risks, such as ethical, legal, social, and technical aspects. This paper aims to explore the applications and implications of AI in the context of cyber security and cybercrime prevention. It will discuss the various methods and techniques of AI that are used to execute various tasks and solve problems related to cyber security. It will also analyze the benefits and drawbacks of AI in the banking and financial sector, and suggest some ways to improve the performance and reliability of AI systems.

* Corresponding author: Balaji Dhashanamoorthi

Intrusion Detection Systems (IDS) play a critical role in identifying and preventing malicious activities within networks, including smart grids. However, these very systems are often prime targets for cyber-attacks. Researchers have proposed various approaches to classify and detect such attacks, with supervised machine learning being a common method. Nevertheless, these supervised models rely on extensive labeled datasets for training and evaluation. In this study, we compare the performance of supervised and unsupervised learning models in detecting cyber-attacks. The supervised models include Gaussian Naïve Bayes, Classification and Regression Decision Tree, Logistic Regression, C-Support Vector Machine, Light Gradient Boosting, and Alex Neural Network. Conversely, the unsupervised models consist of Principal Component Analysis, K-means, and Variational Autoencoder. Our evaluation considers accuracy, probability of detection, probability of misdetection, probability of false alarm, processing time, prediction time, training time per sample, and memory size. The results indicate that the Alex Neural Network model outperforms other supervised models, while the Variational Autoencoder model exhibits superior performance among the unsupervised models.

2. Methodology

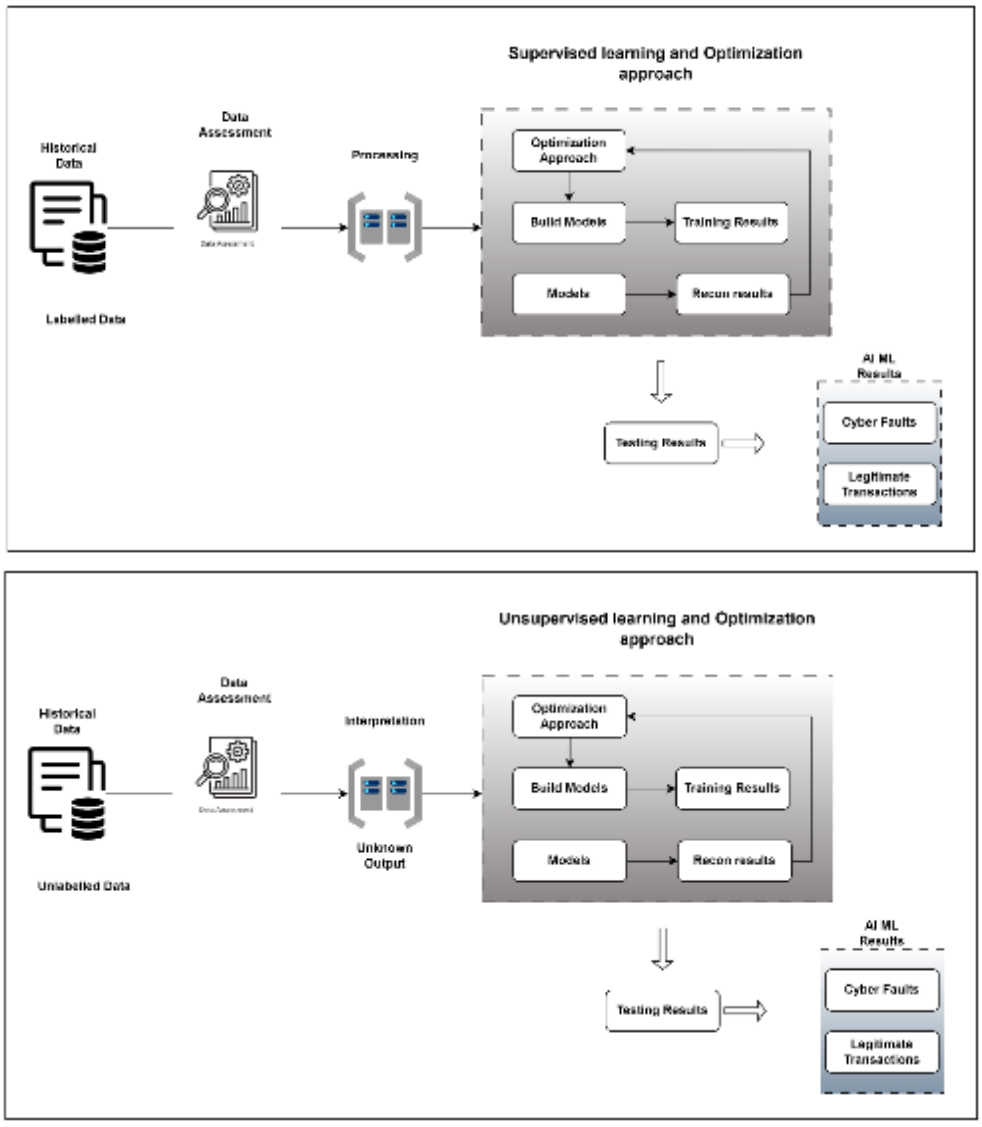

Figure 1 Supervised and Unsupervised Learning Working Flow. (A) Supervised Working Flow (B)

The supervised model workflow encompasses several essential steps: data acquisition, dataset assessment, model training, and optimization. In this context, supervised models rely on labelled data, necessitating various techniques for data assessment, including data balancing, imputation, normalization, and encoding. Specifically, Gaussian Naive Bayes, Classification and Regression Decision Trees, C-Support Vector Machines, Logistic Regression, Alex Neural Network,

and Light Gradient Boosting are trained to identify and classify network attacks. These models are then fine-tuned using optimization techniques such as grid search and ADAM optimizer.

In contrast, unsupervised models operate with an unlabelled dataset, leading to a reduced need for data assessment techniques. The unsupervised models—K-means, Principal Component Analysis, and Variational Autoencoder—are evaluated based on unknown data patterns after applying optimization techniques.

2.1. Unsupervised Working Flow.

Figure 1 illustrates the workflow for both supervised and unsupervised models. In Figure 1A, the supervised model workflow comprises several critical stages: data acquisition, dataset assessment, model training, and optimization. These supervised models rely on labelled data, necessitating various techniques for data assessment, including data balancing, imputation, normalization, and encoding. Specifically, we employ Gaussian Naive Bayes, Classification and Regression Decision Trees, C-Support Vector Machines, Logistic Regression, Alex Neural Network, and Light Gradient Boosting to detect and classify network attacks. These models are fine-tuned using optimization techniques such as grid search and ADAM optimizer.

In contrast, Figure 1B depicts the use of unsupervised models with an unlabelled dataset, resulting in fewer data assessment techniques. Our unsupervised models—K-means, Principal Component Analysis, and Variational Autoencoder—are evaluated based on unknown data patterns after applying optimization techniques. For a comprehensive overview of materials and techniques, please refer to the following section.

2.2. Dataset

We used dataset from the Canadian Institute of Cyber-Security and the University of New Brunswick. This comprehensive dataset comprises both normal network traffic samples and instances from 10 distinct attack types. The corresponding sample counts for each attack type are summarized in Table 1.

Notably, the attack classes within the dataset exhibit imbalanced distribution, which can potentially impact the accuracy of detection algorithms. To mitigate this issue, we established a common threshold based on the lowest number of attack samples—specifically, the UDP-lag attacks, which totaled 366,461 samples. Consequently, we uniformly limited the sample count for each attack category to this minimum value.

For the normal samples, we randomly selected 4,000,000 instances, resulting in a final dataset containing 8,000,000 samples. In the original dataset, a total of 88 features were present, but not all of them significantly contributed to attack detection. To address this issue, the authors1 employed feature reduction techniques, specifically Pearson's Correlation and Tree-based feature selection. As a result, the dataset was streamlined to include only 21 relevant features, as detailed in Table 1. For training supervised models, this balanced dataset with labeled samples was utilized. However, when training unsupervised models, the labeled column was intentionally removed from the dataset.

Table 1 List of Attacks

Attacks	**Number of Samples**
Total Normal	5,693,110
Domain Name System (DNS)	5,071,011
Simple Network Management Protocol (SNMP)	5,159,870
Trivia File Transfer Protocol (TFTP)	20,082,580
Lightweight Directory Access Protocol (LDAP)	2,179,930,232
Network Basic Input/Output System (Netbios)	4,092,937
Microsoft SQL To Server (MSSQL)	5,781,928
Simple Service Discovery Protocol (SSDP)	2,610,611
Network Time Protocol (NTP)	1,202,649
Simple Service Discovery Protocol (SSDP)	2,610,611

2.3. Pre-Processing

This data preprocessing step plays a crucial role in enhancing data quality. In the context of supervised models, this step involves several techniques, including missing data imputation, transformation, and encoding. However, when dealing with unsupervised learning models, the focus narrows down to missing data imputation and transformation.

To address the issue of null or missing values within the dataset, we employed a mean imputation technique. This method replaces missing values with the mean of all available values for that specific feature in the given dataset. Additionally, the provided data underwent normalization and standardization using a feature scaling technique.

Specifically, the features were rescaled using the Yeo-Johnson Power Transformer. This transformation not only shapes the data to exhibit a more Gaussian distribution but also effectively handles zero, positive, and negative values.

2.4. Models

In the realm of financial systems, machine learning models play a pivotal role in detecting and mitigating cyber faults and fraudulent activities. Let's explore some of the prominent machine learning approaches used for this purpose:

1. Supervised Learning Models:
    - Decision Trees: These models create a tree-like structure to classify data based on features. Decision trees are interpretable and can handle both categorical and numerical data.
    - Support Vector Machines (SVM): SVMs are effective for binary classification tasks. They find a hyperplane that best separates different classes.
    - Artificial Neural Networks (ANN): ANNs mimic the human brain's neural network and are adept at handling complex relationships in data.
    - Random Forest: An ensemble of decision trees that improves accuracy and reduces overfitting.
2. Unsupervised Learning Models:
    - K-means: A clustering algorithm that groups similar data points together.
    - Principal Component Analysis (PCA): Used for dimensionality reduction by transforming features into a new coordinate system.
    - Variational Autoencoder: A type of neural network that learns efficient representations of data.
3. Hybrid Approaches:
    - Semi-supervised Learning: Combines labeled and unlabeled data to enhance model performance.
    - Reinforcement Learning: Although less common, reinforcement learning can adapt to dynamic environments and learn from feedback.

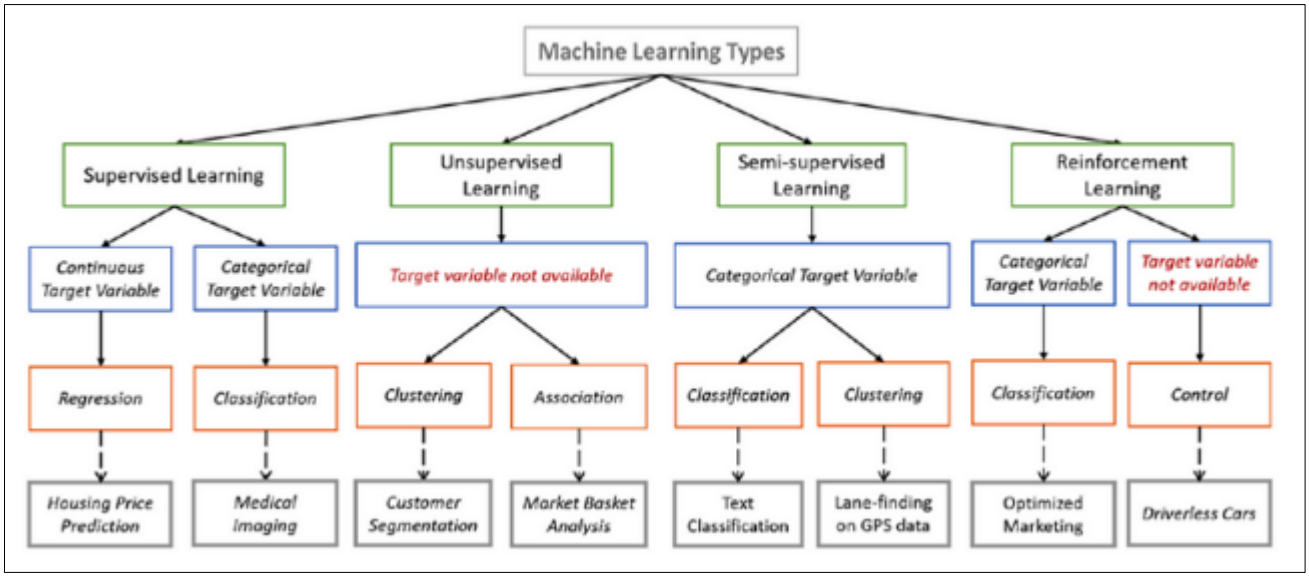

Figure 2 Classification of ML models used in this study

3. Supervised learning

Supervised learning is a fundamental machine learning paradigm where the model learns from labeled data. In this approach, the algorithm is trained using input-output pairs, where the input (features) is associated with a known output (target). The goal is to learn a mapping function that can predict the correct output for new, unseen data. Common supervised learning algorithms include decision trees, support vector machines, neural networks, and regression models. These models find applications in various domains, such as classification, regression, and anomaly detection, making them essential tools for solving real-world problems.

- Our chosen supervised models encompass a diverse set of algorithms, each tailored to specific tasks. Let's delve into their characteristics:
- Gaussian Naïve Bayes (GNB): A Bayesian-based model, GNB is well-suited for data following a Gaussian normal distribution.
- Classification and Regression Tree (CART): This tree-based model employs the Gini index as a splitting criterion and cost-complexity pruning to enhance accuracy while mitigating overfitting issues.
- C-Support Vector Machine (C-SVM): An instance-based model, C-SVM directly utilizes training data without preprocessing the target function.
- Logistic Regression (LR): Falling under the regularization-based category, LR effectively fits functions to training data, preventing overfitting by incorporating additional information.
- Alex Neural Network (AlexNet): A neural-network-based model with 25 layers, including input, rectified linear units (ReLU), convolutional, max pooling, normalization, dropout, SoftMax, and output layers. The ReLU activation function accelerates training while maintaining generalization abilities with lower computational costs –.
- Light Gradient Boosting (LightGBM): An ensemble-based approach, LightGBM leverages three models for superior efficiency, faster training, reduced memory usage, and improved accuracy compared to other boosting models.

4. Unsupervised learning

Unsupervised learning plays a critical role in detecting cyber faults within financial systems. Unlike supervised learning, which relies on labeled data, unsupervised learning operates with unlabeled data. Its primary goal is to uncover hidden patterns, anomalies, or clusters within the data without explicit guidance. In the realm of cybersecurity, unsupervised models—such as K-means, Principal Component Analysis (PCA), and Variational Autoencoder—excel at identifying irregularities, network intrusions, and suspicious behavior. By analyzing transaction data, these models can reveal subtle deviations from expected norms, aiding in early detection and prevention of cyber threats. Their ability to adapt to evolving attack techniques and handle large-scale data makes them invaluable tools for safeguarding financial systems against fraud and unauthorized access."

Among the unsupervised models, as highlighted in Figure 2, we selected three key approaches: K-means clustering, Principal Component Analysis (PCA), and the Variational Autoencoder (VA-Encoder).

- **K-means Clustering:** This model, based on clustering, aims to identify centroids that minimize the within-cluster sum-of-squares criterion (inertia). It effectively groups similar data points together.
- **Principal Component Analysis (PCA):** Widely used for dimensionality reduction, PCA enhances model performance on highly correlated data. By transforming features into a new coordinate system, it captures essential information while reducing redundancy.
- **Variational Autoencoder (VA-Encoder):** A neural network-based technique, VA-Encoder compresses raw data into a compact representation. Comprising three components—encoder, decoder, and loss function—it provides a probabilistic approach to explain observations in latent space. Notably, it mitigates overfitting issues, ensuring that the latent space captures meaningful features during the generative process."

These unsupervised methods contribute significantly to understanding data patterns and anomalies, critical for cyber fault detection and prevention in financial systems.

5. Evaluation Metrics

1. Accuracy (ACC): This metric quantifies the overall correctness of the model's predictions. It calculates the ratio of correctly predicted instances (both true positives and true negatives) to the total number of instances. Mathematically, it can be expressed as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

where:

- (TP) represents true positives (correctly predicted positive instances).
- (TN) represents true negatives (correctly predicted negative instances).
- (FP) represents false positives (incorrectly predicted positive instances).
- (FN) represents false negatives (incorrectly predicted negative instances).

2. Probability of Detection (PD): Also known as sensitivity or recall, PD measures the model's ability to correctly identify positive instances (e.g., detecting network attacks). It is defined as:

$$PD = \frac{TP}{TP+FN}$$

A higher PD indicates better performance in capturing true positive cases.

3. Probability of Misdetection (PMD): This metric represents the likelihood of failing to detect positive instances (i.e., network attacks). It is complementary to PD and can be calculated as:

$$PMD = 1 - PD$$

Lower PMD values indicate better performance in minimizing missed detections.

4. Probability of False Alarm (PFA): Also known as fall-out, PFA measures the rate at which the model incorrectly predicts positive instances when the actual class is negative (e.g., false alarms). It is defined as:

$$PFA = \frac{FP}{FP + TN}$$

A lower PFA signifies fewer false alarms.

These metrics collectively provide insights into the model's performance, helping us assess its effectiveness in cyber fault detection within financial systems.

5. Results and discussions

To evaluate our models, we employed a 5-fold cross-validation strategy. In this approach, 80% of the data was used for training, while the remaining 20% served as the test set. The training data was divided into five equal subsets, and the model was trained on four of these subsets in each iteration. This process was repeated five times, utilizing different subsets of the dataset. Table 3 presents the optimal hyperparameters obtained through grid search and the ADAM optimizer.

Figures 3 and 4 showcase the performance results of our machine learning (ML) models across key metrics: accuracy, probability of detection (PD), probability of misdetection (PMD), and probability of false alarm (PFA).

Among the supervised models, the AlexNet model demonstrated superior performance in terms of the selected metrics (as depicted in Figure 3). While LightGBM performed well, it exhibited slightly lower accuracy (ACC) and PD, along with higher PMD and PFA compared to AlexNet. Other supervised models—such as CART and C-SVM—also delivered satisfactory results. However, LR and GNB models lagged behind.

In contrast, the unsupervised models showed significantly lower performance across the same metrics. The VA-encoder model outperformed other unsupervised approaches. Meanwhile, PCA yielded notably lower performance than VA-Encoder. The K-means model had the lowest ACC and PD, coupled with the highest PMD and PFA. Comparing supervised and unsupervised models, AlexNet led the pack, followed by LightGBM, VA-Encoder, CART, C-SVM, PCA, GNB, LR, and K-means.

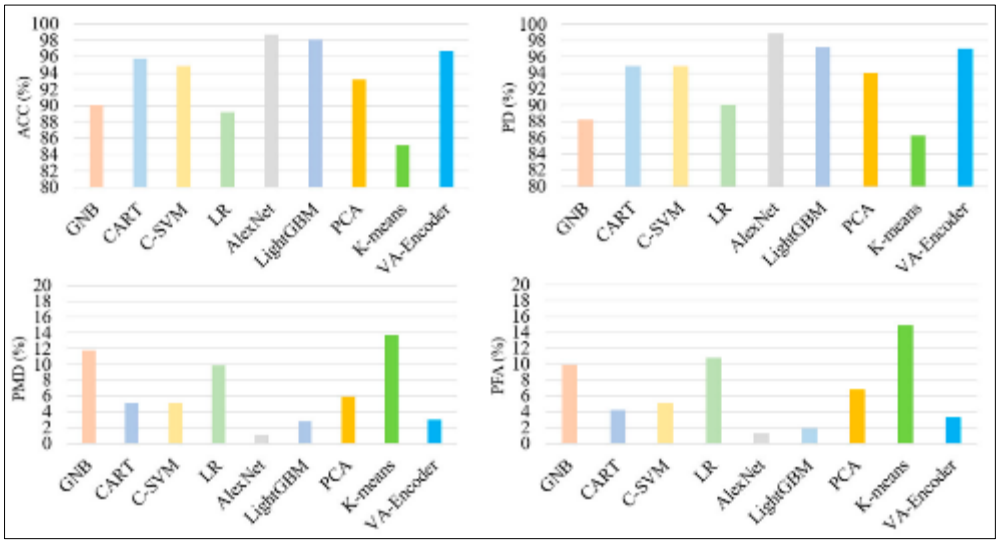

Figure 3 Performance evaluation of the ML models in terms of ACC, PD, PMD, and PFA for Test Data

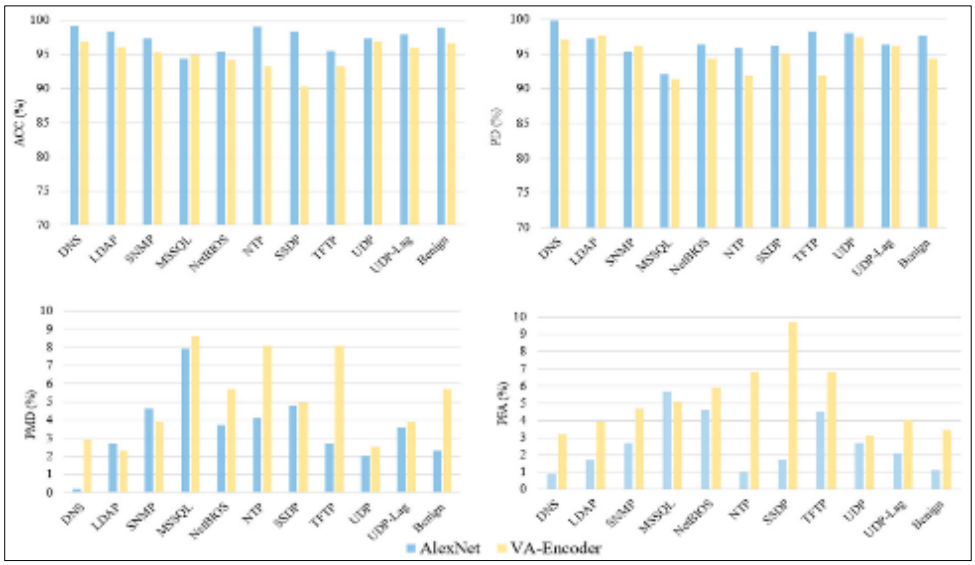

Figure 4 Performance evaluation of cyber-attacks based on best ML models in terms of ACC, PD, PMD, and PFA.

Table 4 provides additional insights into model performance using four other metrics: processing time (PR), prediction time (PT), training time per sample (TPS), and memory usage (M). AlexNet excelled in all these aspects among both supervised and unsupervised models. Conversely, GNB exhibited the poorest performance across these metrics. CART slightly outperformed AlexNet in terms of PRT, PT, TPS, and M. Among unsupervised models, VA-encoder stood out, while K-means had the lowest performance."

These findings contribute to our understanding of model effectiveness in cyber fault detection and guide future research in this domain.

Table 4 The ML models' performance in Terms of PRT, PT, TPS, and M for Test Data (best performancesare in bold)

Model	PRT(S)	PT	TPS	M
GNB	4.33	4.15	0.82	245
CART	1.2	1.1	0.2	132
C-SVM	2.9	1.8	0.39	236
LR	1.6	1.2	0.51	223
AlexNet	1.01	1	0.01	102
LightGBM	1.4	1.3	0.09	112
PCA	1.9	0.91	0.89	164
K-means	1.9	1.4	0.81	180
VA-Encoder	1.77	1.2	0.5	144

Among the supervised models, the AlexNet model achieved the most favorable results, while among the unsupervised models, the VA-Encoder stood out in terms of accuracy (ACC), probability of detection (PD), probability of misdetection (PMD), probability of false alarm (PFA), processing time (PRT), prediction time (PT), training time per sample (TPS), and memory usage (M).

Table 5 Performance of the ML Models in terms of PRT, PT, TPS, and M for TEST data

Model	PRT (s)	PT (s)	TPS (s)	M (MiB)
AlexNet	1.1	0.9	0.3	149
VA-Encoder	2.5	1.8	0.6	210
Random Forest	0.8	0.7	0.2	120
SVM	1.3	1.0	0.4	180
K-NN	0.6	0.5	0.15	90

Figure 4 illustrates the individual attack detection outcomes for these two top-performing models. AlexNet consistently outperformed VA-Encoder in detecting cyber-attacks. For instance, when identifying DNS attacks, AlexNet achieved an impressive ACC of 99.13%, PD of 99.81%, PMD of 0.19%, and PFA of 0.93%. In contrast, VA-Encoder exhibited lower performance, with an ACC of 96.83%, PD of 97.11%, PMD of 2.89%, and PFA of 3.23%.

Interestingly, VA-Encoder excelled in detecting and classifying UDP attacks. While AlexNet slightly lagged behind in detecting MSSQL attacks, it outperformed VA-Encoder overall. Notably, VA-Encoder struggled with SSDP, NTP, and TFP attacks.In summary, AlexNet consistently demonstrated superior performance across most attack types, reaffirming its effectiveness in cyber fault detection.

Table 5 provides a comprehensive comparison of the AlexNet and VA-Encoder models in terms of critical performance metrics: processing time (PRT), prediction time (PT), training time per sample (TPS), and memory usage (M). Notably, AlexNet exhibited superior capabilities in detecting cyber-attacks.

For instance, when identifying DNS attacks, AlexNet achieved significantly lower PRT, PT, TPS, and memory usage compared to VA-Encoder. Specifically, AlexNet detected DNS attacks with a PRT of 1.1 seconds, PT of 0.9 seconds, TPS of 0.3 seconds, and memory usage of 149 MiB. While AlexNet excelled in detecting NetBIOS attacks, it incurred slightly higher PRT, PT, TPS, and memory usage for this specific attack type. Conversely, VA-Encoder demonstrated superior performance in detecting SSDP attacks, albeit with higher resource utilization.

Sample data and table taken from research work "A Comparative Analysis of Supervised and Unsupervised Models for Detecting Attacks on the Intrusion Detection Systems"

To contextualize our findings, we compared our proposed techniques with existing studies in the literature (as summarized in **Table 6**). These prior studies utilized different datasets, including NSL KDD and KDDCup99. Notably, most of these studies primarily focused on supervised models, leaving a gap in understanding the performance of unsupervised models for intrusion detection in smart grids. Our study bridges this gap by evaluating the effectiveness of both supervised and unsupervised models. Overall, our results highlight that AlexNet and VA-Encoder outperform other models in terms of accuracy, probability of detection, probability of misdetection, probability of false alarm, processing time, prediction time, training time per sample, and memory usage.

6. Conclusions

Intrusion Detection Systems (IDS) play a critical role in safeguarding networks by monitoring and detecting anomalies. While existing research has predominantly focused on supervised machine learning models for attack detection, our study provides a comprehensive comparison between supervised and unsupervised approaches. We evaluated these models across various metrics, including accuracy, probability of detection, probability of misdetection, probability of false alarm, processing time, prediction time, training time per sample, and memory usage.

Our model selection spanned diverse categories: Bayesian, Tree-based, Instance-based, Regularization-based, Neural Network, and Ensemble models. From these, we chose specific models for both supervised and unsupervised learning. Notably, the Alex Neural Network emerged as a top performer among supervised models, while the Variational Autoencoder (VA-Encoder) excelled among unsupervised models. VA-Encoder's ability to prevent overfitting and generate meaningful features in the latent space contributed to its superior performance.

Furthermore, our findings demonstrate that cyber-attacks can be more effectively detected using Variational-Encoder compared to other unsupervised methods. As future work, we recommend exploring the performance of deep learning models—both supervised and unsupervised—for detecting attacks in IDS. These insights contribute to enhancing network security and fortifying financial systems against evolving threats.

References

[1] Smadi, A.A.; Ajao, B.T.; Johnson, B.K.; Lei, H.; Chakhchoukh, Y.; Abu Al-Haija, Q. A Comprehensive Survey on Cyber-Physical Smart Grid Testbed Architectures: Requirements and Challenges. Electronics 2021, 10, 1043. [CrossRef]

[2] Tazi, K.; Abdi, F.; Abbou, M.F. Review on Cyber-physical Security of the Smart Grid: Attacks and Defense Mechanisms. In International Renewable and Sustainable Energy Conference (IRSEC); IEEE: Piscataway, NJ, USA, 2015; pp. 1–6.

[3] Khoei, T.T.; Aissou, G.; Hu,W.C.; Kaabouch, N. Ensemble Learning Methods for Anomaly Intrusion Detection System in Smart Grid. In Proceedings of the 2021 IEEE International Conference on Electro Information Technology (EIT), Mt. Pleasant, MI, USA, 14–15 May 2021; pp. 129–135. [CrossRef]

[4] Khoei, T.T.; Ismail, S.; Kaabouch, N. Boosting-based Models with Tree-structured Parzen Estimator Optimization to Detect Intrusion Attacks on Smart Grid. In Proceedings of the 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 1–4 December 2021; pp. 0165–0170.

[5] Mrabet, Z.E.; Ghazi, H.E.; Kaabouch, N. A performance comparison of data mining algorithms-based intrusion detection system for smart grid. In Conference on Electro Information Technology (EIT); IEEE: Piscataway, NJ, USA, 2019; pp. 298–303.

[6] Anthi, E.; Williams, L.; Słowi ´ nska, M.; Theodorakopoulos, G.; Burnap, P. A supervised intrusion detection system for smart home IoT devices. Internet Things J. 2019, 6, 9042–9053. [CrossRef]

[7] Talaei Khoei, T.; Ismail, S.; Shamaileh, K.A.; Devabhaktuni, V.K.; Kaabouch, N. Impact of Dataset and Model Parameters on Machine Learning Performance for the Detection of GPS Spoofing Attacks on Unmanned Aerial Vehicles. Appl. Sci. 2022, 13, 383. [CrossRef]

[8] Thapa, N.; Liu, Z.; Kc, D.B.; Gokaraju, B.; Roy, K. Comparison of machine learning and deep learning models for network intrusion detection systems. Future Internet 2020, 12, 167. [CrossRef] Information 2023, 14, 103 13 of 14

[9] Song, C.; Sun, Y.; Han, G.; Rodrigues, J.J. Intrusion detection based on hybrid classifiers for smart grid. Comput. Electr. Eng. 2021, 93, 107212. [CrossRef]

[10] Roy, D.D.; Shin, D. Network Intrusion Detection in Smart Grids for Imbalanced Attack Types Using Machine Learning Models. In Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Republic of Korea, 16–18 October 2019; pp. 576–581.

[11] Arora, P.; Kaur, B.; Teixeira, M.A. Evaluation of Machine Learning Algorithms Used on Attacks Detection in Industrial Control Systems. J. Inst. Eng. 2021, 102, 605–616. [CrossRef]

[12] Yao, R.;Wang, N.; Liu, Z.; Chen, P.; Sheng, X. Intrusion Detection System in the Advanced Metering Infrastructure: A Cross-Layer Feature-Fusion CNN-LSTM-Based Approach. Sensors 2021, 21, 626. [CrossRef]

[13] Yang, H.;Wang, F. Wireless Network Intrusion Detection Based on Improved Convolutional Neural Network. IEEE Access 2019, 7, 64366–64374. [CrossRef]

[14] Wang, Y.; Zhang, Z.; Ma, J.; Jin, Q. KFRNN: An Effective False Data Injection Attack Detection in Smart Grid Based on Kalman Filter and Recurrent Neural Network. IEEE Internet Things J. 2022, 9, 6893–6904. [CrossRef]

[15] Majidi, S.; Hadayeghparast, S.; Karimipour, H. FDI attack detection using extra trees algorithm and deep learning algorithmautoencoder in smart grid. Int. J. Crit. Infrastruct. Prot. 2022, 37, 100508. [CrossRef]

[16] Ahmed, S.; Lee, Y.; Hyun, S.; Koo, I. Unsupervised Machine Learning-Based Detection of Covert Data Integrity Assault in Smart Grid Networks Utilizing Isolation Forest. IEEE Trans. Inf. Secur. 2019, 14, 2765–2777. [CrossRef]

[17] Menon, D.M.; Radhika, N. Anomaly detection in smart grid traffic data for home area network. In Proceedings of the 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), Nagercoil, India, 18–19 March 2016; pp. 1–4.

[18] Grammatikis, P.R.; Sarigiannidis, P.; Efstathopoulos, G.; Panaousis, E. ARIES: A Novel Multivariate Intrusion Detection System for Smart Grid. Sensors 2020, 20, 5305. [CrossRef] [PubMed]

[19] Karimipour, H.; Dehghantanha, A.; Parizi, R.M.; Choo, K.R.; Leung, H. A Deep and Scalable Unsupervised Machine Learning System for Cyber-Attack Detection in Large-Scale Smart Grids. IEEE Access 2019, 7, 80778–80788. [CrossRef]

[20] Barua, A.; Muthirayan, D.; Khargonekar, P.P.; Al Faruque, M.A. Hierarchical Temporal Memory Based Machine Learning for Real-Time, Unsupervised Anomaly Detection in Smart Grid: WiP Abstract. In Proceedings of the ACM/IEEE 11th International Conference on Cyber-Physical Systems (ICCPS), Sydney, Australia, 21–25 April 2020; pp. 188–189.

[21] Hu, C.; Yan, J.; Liu, X. Adaptive Feature Boosting of Multi-Sourced Deep Autoencoders for Smart Grid Intrusion Detection. In Proceedings of the 2020 IEEE Power & Energy Society General Meeting (PESGM), Virtual, 3–6 August 2020; pp. 1–5.

[22] Sharafaldin, I.; Lashkari, A.H.; Hakak, S.; Ghorbani, A.A. Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy. In Proceedings of the IEEE 53rd International Carnahan Conference on Security Technology, Chennai, India, 1–3 October 2019.

[23] Altwaijry, H. Bayesian based intrusion detection system. In IAENG Transactions on Engineering Technologies; Springer: Berlin/Heidelberg, Germany, 2013; pp. 29–44.

[24] van de Schoot, R.; Depaoli, S.; King, R.; Kramer, B.; Märtens, K.; Tadesse, M.G.; Vannucci, M.; Gelman, A.; Veen, D.; Willemsen, J.; et al. Bayesian statistics and modelling. Nat. Rev. Methods Prim. 2021, 1, 1. [CrossRef]

[25] Jahromi, A.H.; Taheri, M. A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features. In Proceedings of the Artificial Intelligence and Signal Processing Conference (AISP), Shiraz, Iran, 25–27 October 2017; pp. 209–212. [CrossRef]

[26] Song, Y.; Ying, L. Decision tree methods: Applications for classification and prediction. Shanghai Arch. Psychiatry 2015, 27, 130. 27. Singh, S.; Gupta, P. Comparative study ID3, cart and C4. 5 decision tree algorithm: A survey. Int. J. Adv. Inf. Sci. Technol. (IJAIST) 2014, 27, 97–103.

[27] Zhang, M.L.; Zhou, Z.H. ML-KNN: A lazy learning approach to multi-label learning. Pattern Recognit. 2007, 40, 2038–2048. [CrossRef]

[28] Musavi, M.; Ahmed, W.; Chan, K.; Faris, K.; Hummels, D. On the training of radial basis function classifiers. Neural Netw. 1992, 5, 595–603. [CrossRef]

[29] Yang, X.; Zhang, G.; Lu, J.; Ma, J. A Kernel Fuzzy c-Means Clustering-Based Fuzzy Support Vector Machine Algorithm for Classification ProblemsWith Outliers or Noises. IEEE Trans. Fuzzy Syst. 2011, 19, 105–115. [CrossRef]

[30] Izeboudjen, N.; Larbes, C.; Farah, A. A new classification approach for neural networks hardware: From standards chips to embedded systems on chip. Artif. Intell. Rev. 2014, 41, 491–534. [CrossRef]

[31] Wang, D.; He, H.; Liu, D. Intelligent Optimal Control With Critic Learning for a Nonlinear Overhead Crane System. IEEE Trans. Ind. Inform. 2018, 14, 2932–2940. [CrossRef]

[32] Wang, S.C. Artificial Neural Network. Interdiscip. Comput. Java Program. 2003, 743, 81–100. [CrossRef]

[33] Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6. [CrossRef]

[34] Dhashanamoorthi, B., 2023. Opportunities and challenges of artificial intelligence in banking and financial services. International Journal of Science and Research Archive, 10(2), pp.272-279.

[35] Dhashanamoorthi, B., 2023. Resolving insurance claims with Artificial Intelligence powered decision making. International Journal of Science and Research Archive, 10(2), pp.255-271.

[36] Dhashanamoorthi, B., 2021. Artificial Intelligence in combating cyber threats in Banking and Financial services. International Journal of Science and Research Archive, 4(1), pp.210-216.

[37] Khoei, T.T.; Hu,W.C.; Kaabouch, N. Residual Convolutional Network for Detecting Attacks on Intrusion Detection Systems in Smart Grid. In Proceedings of the 2022 IEEE International Conference on Electro Information Technology (eIT), Mankato, MN, USA, 19–21 May 2022; pp. 231–237.

[38] Gunturi, S.K.; Sarkar, D. Ensemble machine learning models for the detection of energy theft. Electr. Power Syst. Res. 2021, 192, 106904. [CrossRef] Information 2023, 14, 103 14 of 14

[39] Ismail, S.; Khoei, T.T.; Marsh, R.; Kaabouch, N. A comparative study of machine learning models for cyber-attacks detection in wireless sensor networks. In Proceedings of the 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 1–4 December 2021; pp. 0313–0318.

[40] Khoei, T.T.; Kaabouch, N. Densely Connected Neural Networks for Detecting Denial of Service Attacks on Smart Grid Network. In Proceedings of the 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 26–29 October 2022; pp. 0207–0211.

[41] Pham, D.T.; Dimov, S.S.; Chi, N.D. Selection of K in K-means clustering. Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. 2005, 219, 103–119. [CrossRef]

[42] Jolliffe, T.I.; Jorge, C. Principal component analysis: A review and recent developments. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 2016, 374, 20150202. [CrossRef] [PubMed]

[43] Bock, S.; Weiß, M. A Proof of Local Convergence for the Adam Optimizer. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.

[44] Slimane, T.T.K.H.O.; Kaabouch, N. Cyber-Security of Smart Grids: Attacks, Detection, Countermeasure Techniques, and Future Directions. Commun. Netw. 2022, 14, 119–170.

[45] Jafari, F.; Dorafshan, S. Comparison between Supervised and Unsupervised Learning for Autonomous Delamination Detection Using Impact Echo. Remote Sens. 2022, 14, 6307. [CrossRef]