International Journal of IT & Knowledge Management (ISSN: 0973-4414)
Volume 15 •Issue 2 pp. 96-106 Jan-June 2022 www.csjournals.com

Advanced IDS using Deep Learning, Feature Selection, and Optimization Techniques

[1]Navroop Kaur, [2]Meenakshi Bansal, [3]Sukhwinder Singh Sran

[1]Research Scholar (Ph.D.), Punjabi University, Patiala, Punjab, India.

[2]Associate Professor, CSE, Yadavindra Department of Engineering, Talwandi Sabo, India.

[3]Assistant professor, Department of Computer Science & Engineering, Punjabi University, Patiala, India.

knavroop7488@gmail.com, ermeenu10@gmail.com, sukhwinder.sran@gmail.com

Abstract:-Intrusion Detection Systems (IDS) play a vital role in safeguarding against the growing spectrum of cyber security threats, particularly in complex environments such as the Internet of Things (IoT), cloud computing, and industrial networks. This study presents a comprehensive evaluation of current state-of-the-art IDS methodologies, with a focus on Deep Learning (DL) techniques and advanced feature engineering strategies. It further explores the effectiveness of optimization models like Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Explainable Artificial Intelligence (XAI) in enhancing detection accuracy, computational efficiency, and system interoperability.

Keywords: - Intrusion Detection Systems (IDS); machine learning; network security; Intrusion Prevention Systems (IPS); deep learning algorithms

1. Introduction: - Emerging technologies such as block chain and quantum computing are also examined for their potential to strengthen data privacy, scalability, and resilience in decentralized and resource-constrained infrastructures [8,10]. The study identifies major challenges in the IDS domain, including real-time anomaly detection, adversarial robustness, and the handling of imbalanced datasets. Addressing these challenges is essential to guide future research aimed at developing lightweight, adaptive, and ethically responsible IDS frameworks capable of real-time operation in dynamic and heterogeneous environments.

By critically synthesizing current IDS approaches, research gaps, and innovative cyber security strategies, this work serves as a valuable reference for academics, researchers, and industry professionals seeking to advance the field of intrusion detection.

To address the identified challenges in intrusion detection, this paper introduces an Interpretable and Robust Intrusion Detection System (IR-IDS) designed for fine-grained classification of known attacks and effective detection of unknown anomalies [7]. The proposed IR-IDS framework is composed of two primary stages: causal feature selection and robust anomaly classification. These components operate independently, allowing seamless integration with various classification and feature selection algorithms, thus offering high flexibility and adaptability.

The causal feature selection module employs conditional testing to assess the causal impact of features on the target variable. By leveraging the principles of conditional independence and the Markov blanket, the method efficiently identifies an optimal subset of features, ensuring both accuracy and computational efficiency in the selection process.

In the next stage, the selected features are utilized to optimize a decision tree classifier using Shapley values, enhancing the model's interpretability and enabling effective pre-classification of network traffic. The outputs are then fed into a hybrid model combining Kolmogorov–Arnold Networks (KAN) with a Conditional Variational Autoencoder (CVAE)—referred to as KAN–CVAE. This model uses reconstruction errors to detect anomalies linked to previously unseen attacks. The integration of KAN within the CVAE architecture significantly improves the interpretability and performance of the system in complex detection tasks.

2. Key contributions of this work include [2,4]:

1. An efficient Markov blanket search algorithm grounded in causal feature selection, enabling accurate identification of the most relevant features for target prediction.
2. A novel conditional testing method based on propensity scores, which uncovers causal dependencies between features and target variables while maintaining interpretability and mitigating high-dimensional data challenges.
3. An interpretable classification model, Shapley Tree, which employs Shapley values to enhance pre-classification of attack traffic, supporting fine-grained and transparent decision-making.

4. The development of KAN–CVAE, an interpretable encoding model that detects unknown attack anomalies through reconstruction error analysis, further strengthened by Extreme Value Theory (EVT) to improve anomaly detection precision.

2.1 A **Markov Blanket Search Algorithm** in causal feature selection identifies the **minimum set of features** that makes the target variable **conditionally independent** of all other features. This ensures **optimal predictive accuracy** while eliminating irrelevant or redundant data.

Concept	Explanation
Markov Blanket (MB)	The smallest set of variables (features) that makes the target variable conditionally independent of all others in the dataset.
Causal Feature Selection	Selects features based on cause-effect relationships, not just statistical correlation.
Algorithm Efficiency	Uses search heuristics (e.g., greedy, forward-backward) to reduce computation time.
Search Process	Iteratively adds/removes features based on relevance and redundancy with respect to the target.
Relevance Criterion	Feature provides unique predictive power for the target.
Redundancy Check	Removes features that do not add new information (already captured by others).
Outcome	A compact, interpretable, and predictive feature set (the Markov Blanket of the target).

Table 1 Markov Blanket Search Algorithm

Process Flow 1:- Optimal predictive accuracy

Raw Dataset

Preprocessing → Identify Target Variable

Initialize MB Set = ∅

Evaluate All Features:

- Measure Causal Relevance to Target

- Add Feature if Improves Predictive Independence

Check Redundancy:

- Remove Redundant Features from MB Set

Output: Optimal Markov Blanket (Relevant + Non-redundant Features)

2.2 This method uses **propensity scores** the probability of receiving a treatment (or having a feature) given other variables to perform **conditional independence tests**. It helps identify **causal relationships** between features and a target while addressing problems common in high-dimensional datasets (like over fitting, noise, and irrelevant variables) [6,8]. The approach remains interpretable because it leverages score-based matching or stratification instead of complex black-box models.

Component	Explanation
Propensity Score	Probability of observing a feature given other covariates (used to balance data like in experiments).
Conditional Testing	Statistical method to test if the relationship between a feature and target holds **given** other variables.
Causal Dependency	A cause-effect relationship between a feature and the target (beyond correlation).
Interpretability	Maintains clear logic (e.g., logistic regression for propensity) rather than complex models.

Component	Explanation
High-Dimensional Challenge	Addresses issues from large feature sets (e.g., noise, multicollinearity) by reducing spurious links.
Matching/Stratification	Divides samples into similar groups based on scores to control confounding.

Table 2 Conditional Independence Test

Process flow 2: (Conditional Test) using Python: -

Step 1: Load Dataset

Step 2: Select Target Variable and Covariates

Step 3: For Each Feature Xi:

├—> Estimate Propensity Score P (Xi | other features)

├—> Stratify or Match samples based on score

├—> Perform Conditional Independence Test:

│ - Is Target ⊥ Xi | P (Xi | others)?

└—> Record Result (Dependent or Independent)

Step 4: Aggregate All Causal Features (Xi where dependence found)

Step 5: Output Interpretable Causal Feature Set

2.3 The **Shapley Tree** is an interpretable classification model that integrates **Shapley values** a game theory based feature attribution method into a **decision tree classifier** to enhance the **pre-classification of attack traffic**. This model not only predicts whether network traffic is

malicious or benign but also **explains the contribution of each feature** to that decision in a **transparent** and **fine-grained** manner [12,15].

- **Shapley Values**: Quantify the contribution of each feature to a specific prediction, based on cooperative game theory.
- **Pre-classification**: Enables early detection and flagging of potential attacks before final decisions.
- **Interpretable**: Each decision can be broken down into feature-level contributions, making the model auditable and explainable.
- **Fine-grained Analysis**: Offers per-instance insight, showing which features influenced the decision and how.
- **Suitable for Security Applications**: Especially valuable in cyber security contexts, where **explain ability** is essential for trust and regulatory compliance.

Component	Explanation
Shapley Tree	A decision tree model enhanced with Shapley values to explain predictions.
Shapley Values	Game theory-based scores showing how much each feature contributes to a prediction.
Pre-classification of Attack Traffic	Detects and flags potentially malicious traffic early, before final classification.
Fine-Grained Decisions	Provides detailed insights into why a specific prediction (attack/benign) was made.
Transparent Interpretability	Explains each decision path with feature impact, making it easy to audit and understand.

Table 3 Component with simple Shapley Tree model

Process Flow 3: Shapley Tree model:-

1. **Input**: Network traffic data is collected.
2. **Feature Extraction**: Key features (e.g., packet size, protocol type) are derived.
3. **Model Training**: A decision tree classifier is trained on labeled traffic.
4. **Shapley Value Calculation**: For each prediction, the contribution of each feature is computed.
5. **Pre-classification**: Early detection of attack behavior is flagged.
6. **Explanation Generation**: Shapley values are used to explain *why* the model predicted attack or benign.
7. **Outcome**: Traffic is classified and the decision is made interpretable for auditing or security actions.

2.4 **KAN–CVAE** (Kernel Attention Network–Conditional Variational Auto encoder) is a hybrid **deep learning model** designed to detect **unknown (zero-day) attack anomalies** in cyber security systems. It leverages the **Conditional Variational Auto encoder (CVAE)** for learning compact [14], meaningful latent representations of normal behavior and uses **reconstruction error analysis** to detect deviations that indicate anomalies.

To enhance **precision in anomaly detection**, especially in **extreme outlier cases**, the model incorporates **Extreme Value Theory (EVT)**, a statistical method that models the tail behavior of reconstruction errors—those most likely to indicate rare or severe attacks.

Component	Purpose
CVAE (Conditional VAE)	Encodes inputs into a latent space conditioned on known context (e.g., class/type) to better reconstruct normal traffic.
KAN (Kernel Attention Network)	Enhances the encoder/decoder with attention mechanisms to focus on critical features.
Reconstruction Error	Measures the difference between input and reconstructed output. High error suggests an anomaly.
EVT (Extreme Value Theory)	Models the tail of reconstruction errors to statistically identify rare and significant anomalies.

Component	Purpose
Interpretability	The attention layers and error-based decision boundary help trace why a sample was flagged as an attack.

Table 4: KAN-CVAE Network purpose

Model / Method	Purpose	Interpretability	Application Area	Complexity
1. Markov Blanket Search (Causal FS)	Select the most relevant causal features for prediction	High (causal and minimal feature set)	Feature selection for any predictive model	Medium (search over conditional dependencies)
2. Propensity Score Conditional Testing	Test causal dependencies between features and target	Moderate to High (uses matching and statistics)	Causal inference in high-dimensional data	High (propensity estimation + testing)
3. Shapley Tree Classifier	Classify and interpret attack traffic using Shapley values	High (feature attribution with Shapley values)	Intrusion detection and explainable classification	Medium (tree + Shapley values)
4. KAN–CVAE with EVT	Detect unknown anomalies via reconstruction + EVT	Moderate (indirect via reconstruction and EVT scores)	Anomaly detection in cyber security (zero-day threats)	High (deep learning + statistical EVT)

Table 5. Comparison between 04 model/methods

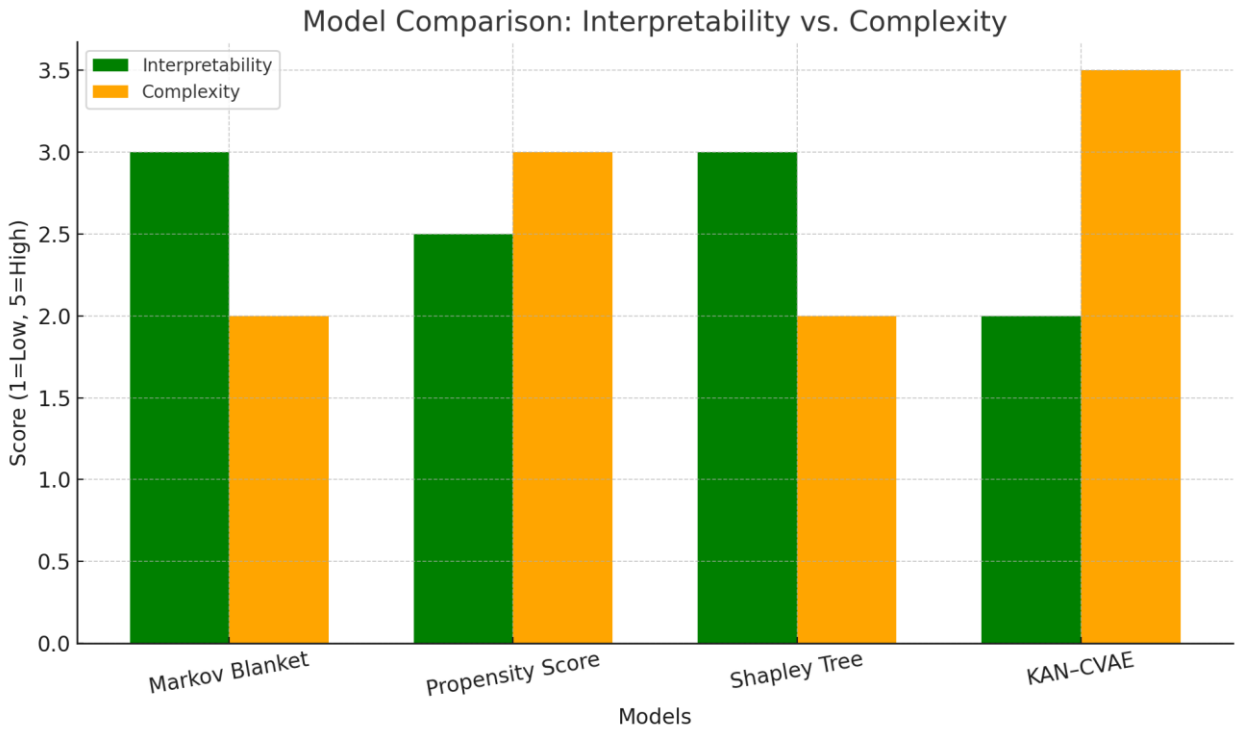

Chart 1: Model Comparison Interpretability V/S Complexity using 04 models

4. Conclusion:-This comprehensive and interpretable framework presents a significant advancement in IDS design, contributing both methodological innovations and practical insights for improving detection accuracy and robustness in real-world cyber security applications.

Firstly, the concept of Intrusion Detection Systems was presented. There are three main types of IDS: Network Intrusion Detection System, Host Intrusion Detection System, and a Hybrid Intrusion Detection System. In addition, each type of IDS can either detect attacks by using a recorded signature or by comparing the behavior of the network with a baseline of the normal traffic or both. Then, the different metrics used to assess Intrusion Detection System by various researchers are presented. The most important metrics are the Accuracy, the Detection Rate (Recall) and the F-Measure.

A general overview of what is machine learning, and a global taxonomy is also discussed. There are three types of machine learning techniques: Supervised, Semi-supervised and Unsupervised.

Most of the machine learning techniques studied fall into one of these categories. A comprehensive review of recently published papers using machine learning for IDS was also discussed.

Based on this study, the recent trends show that deep learning methods are more and more used to detect attacks. However, this increases the complexity of the models which requires more computing resources. In addition, it was shown that more and more solutions are using feature extraction with Auto-Encoder as one of the techniques used.

Reference:-

1. Anderson, P. Computer Security Threat Monitoring and Surveillance. 1980.: https://csrc.nist.gov/csrc/media/ publications/conference-paper/1998/10/08/proceedings-of- the-21st- nissc-1998/documents/early- cs-papers/ande80.pdf (accessed on 19 May 2022).
2. ThreatStack. The History of Intrusion Detection Systems (IDS)—Part 1. Available online: https://www.threatstack.com/blog/ the-history-of-intrusion-detection-systems-ids-part-1 (accessed on 19 May 2022).
3. Checkpoint. What Is an Intrusion Detection System? Available online: https://www.checkpoint.com/cyber- hub/network-security/what-is-an-intrusion-detection-system-ids/ (accessed on 19 May 2022).
4. Sabahi, F.; Movaghar, A. Intrusion Detection: A Survey. In Proceedings of the 2008 Third International Conference on Systems and Networks Communications, Sliema, Malta, 26–31 October 2008; pp. 23–26. [CrossRef]
5. IBM Cloud Education. Machine Learning.:https://www.ibm.com/cloud/learn/machine-learning (accessed on 19 May 2022).
6. Zhang, L., et al. "A Comparative Study on Feature Selection in Intrusion Detection System Using Filter Methods." Journal of Network and Computer Applications.
7. Liu, H., et al. "Feature Selection with Data Cleaning for Cyber-Attack Detection in IoT Systems." IEEE Internet of Things Journal.
8. Jain, A., et al. "Optimization of Intrusion Detection Systems using Genetic Algorithms." Expert Systems with Applications.

9. Liu, D., et al. "An Intrusion Detection Model Based on Improved Genetic Algorithm and SVM." Applied Soft Computing.
10. Deb, K. "Optimization for Engineering Design: Algorithms and Examples." Prentice Hall.
11. Gao, W., et al. "A Novel Deep Learning Based Intrusion Detection System for Industrial Control Systems." IEEE Transactions on Industrial Informatics.
12. Sood, K., et al. "Deep Learning for Anomaly Detection: A Survey." ACM Computing Surveys.
13. Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009).A detailed analysis of the KDD CUP 99 data set. Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications.
14. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700.
15. Sangkatsanee, P., Wattanapongsakorn, N., & Charnsripinyo, C. (2011**).** Practical real-time intrusion detection using machine learning approaches. Computer Communications, 34(18), 2227-2235.
16. An efficient feature selection method for intrusion detection system based on SVM and improved binary differential evolution algorithm.Soft Computing, 24, 12505–12518.