



# Enhancing Robustness of Machine Learning based Intrusion Detection Systems under Distribution Shift through Stability-Constrained Learning

S.Valarmathi, S.Mugambigai

Department of Computer Science and Engineering, Knowledge Institute of Technology (Autonomous), Salem,  
Affiliated to Anna University, Chennai, Tamil Nadu, India

**Publication History:** Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03.2026; Published: 28.03.2026.

**ABSTRACT:** Machine learning is widely used in intrusion detection systems to identify malicious network activities, but in practical environments, the data used during training often differs from the data encountered after deployment because network traffic changes over time. As a result, models that perform well during training may behave inconsistently in real conditions, especially when small variations in input lead to noticeable changes in prediction. This work looks at intrusion detection from the perspective of prediction stability under such changing data conditions by introducing a stability constraint during training that limits how much the model output can vary when the input is slightly modified. Along with classification accuracy, the behavior of the model is examined using measures such as prediction sensitivity, output variance, and parameter magnitude so that reliability can be understood more clearly rather than relying on accuracy alone. The approach is evaluated using benchmark intrusion detection data with controlled variations that represent changes in network traffic, and the observations show that the model produces more consistent outputs and is less affected by small input changes while maintaining a similar level of detection performance, indicating that incorporating stability into the learning process can improve the reliability of intrusion detection systems in dynamic environments.

**KEYWORDS:** Intrusion Detection Systems- Distribution Shift - Stability-Constrained Learning - Network Security- Robust Machine Learning.

## I. INTRODUCTION

Computer networks play an important role in supporting communication, data exchange, and many critical services across different sectors. As their usage continues to grow, concerns related to security have also increased. Network-based attacks such as denial-of-service, malware, and unauthorized access are becoming more frequent and, in many cases, more difficult to detect. This has made intrusion detection an essential component of modern network security.

Intrusion Detection Systems are designed to monitor network traffic and identify behavior that deviates from normal patterns. Traditional approaches are mainly based on signature matching, where known attack patterns are compared with incoming data. Although these methods are effective for previously identified threats, they are limited when dealing with new or evolving attacks. To address this limitation, machine learning techniques have been widely adopted, as they can learn patterns from data and identify anomalies without relying entirely on predefined rules.

Despite these advantages, several practical challenges remain. A key issue is that many machine learning models are developed under the assumption that training data and operational data follow similar distributions. In real-world environments, this assumption rarely holds. Network traffic changes over time due to variations in user behavior, system updates, and the emergence of new attack strategies. When such changes occur, the performance of trained models may degrade, and their predictions may become less reliable.

Another important concern is the stability of model predictions. In some situations, small changes in input features can lead to noticeable differences in the output. In the context of intrusion detection, such behavior can result in inconsistent decisions, increasing the likelihood of false alarms or missed detections. Therefore, it is not sufficient to



evaluate models based only on accuracy; it is also necessary to examine how stable their predictions remain under small variations in input data.

Existing research has explored techniques such as domain adaptation, transfer learning, and robust optimization to handle differences between training and testing data. While these methods provide useful improvements, they often require additional information about the target domain or involve complex training procedures. In many practical scenarios, such information may not be available in advance, making these approaches difficult to apply directly. In this work, intrusion detection is examined from a different perspective by focusing on prediction stability under changing data conditions. A stability-constrained learning framework is introduced, where the training process is designed to reduce unnecessary changes in model output when the input is slightly modified. Instead of adapting the model to a specific target domain, the approach studies how the model itself responds to controlled variations in input data.

To evaluate this behavior, small perturbations are introduced to simulate realistic changes in network traffic. The model is then assessed not only in terms of classification accuracy but also based on how consistent its predictions remain under these variations. This allows a more detailed understanding of model reliability in dynamic environments. The results show that incorporating stability constraints during training leads to more consistent prediction behavior without significantly affecting detection performance. This suggests that focusing on stability, along with accuracy, can help improve the practical usefulness of intrusion detection systems in real-world settings where data conditions are not fixed.

## II. RELATED WORK

Handling changes in data distribution has been widely studied in machine learning, especially in the context of domain adaptation and domain generalization. Existing approaches mainly focus on reducing the gap between training and testing data by learning domain-invariant representations or aligning feature distributions across domains. Several studies have explored feature alignment and transfer learning strategies to improve model performance under distribution shift [1–5,16,17]. More advanced methods, including adversarial domain adaptation and cycle-consistent learning, attempt to minimize discrepancies between source and target domains through complex optimization procedures [18,19]. While these approaches have shown promising results, they often rely on access to target domain data during training or require careful tuning of multiple components, which may limit their applicability in practical intrusion detection scenarios.

Another line of research focuses on improving robustness through distributionally robust optimization and invariant learning. These methods aim to ensure that models perform consistently under worst-case or unseen data conditions [6–10]. Techniques such as adversarial training and invariant risk minimization attempt to reduce sensitivity to distribution changes by enforcing stability across environments. Although these approaches provide strong theoretical foundations, they are often computationally intensive and are not always designed with application-specific constraints in mind. In particular, their direct application to intrusion detection systems remains limited.

The concept of stability in machine learning has also been studied from a theoretical perspective. Earlier work has shown that stable learning algorithms tend to generalize better and produce more reliable predictions [11,12]. Other studies have examined related aspects such as probability calibration and stability selection to improve the reliability of model outputs [13–15]. However, most of these works focus on general learning behavior and do not explicitly address how prediction stability can be evaluated and controlled in the presence of distribution shift, especially in security-related applications.

In the area of intrusion detection, machine learning and deep learning techniques have been widely adopted to improve detection accuracy and automate threat identification [20–22,25,26]. Several surveys have highlighted the effectiveness of these approaches as well as the availability of benchmark datasets such as KDD Cup 1999 [23,27–30]. Despite these advancements, a common limitation is that many intrusion detection models are developed and evaluated under the assumption that training and testing data follow similar distributions. As pointed out in prior work, this assumption does not always hold in real-world environments, where network traffic patterns can evolve over time [24]. As a result, models that perform well during training may experience degraded performance and inconsistent behavior after deployment.



Overall, while existing studies have addressed distribution shift, robustness, and intrusion detection from different perspectives, limited attention has been given to the stability of model predictions under changing input conditions. Most approaches either focus on adapting to new domains or improving accuracy, without explicitly examining how sensitive model outputs are to small variations in input data. In contrast, the present work treats stability as an integral part of the learning process and evaluates model behavior under controlled perturbations. By focusing on both predictive performance and consistency, the proposed approach aims to provide a more practical and reliable solution for intrusion detection in dynamic network environments.

### III. METHODOLOGY

This section presents the methodology developed to improve the reliability of intrusion detection models under changing data conditions. In real network environments, traffic patterns evolve, which can reduce model effectiveness after deployment. The proposed approach is designed not only to learn from available data but also to examine and control how model predictions respond to small variations in input. By combining controlled training with structured evaluation, the methodology provides a way to study both predictive performance and prediction stability under distribution shift.

The dataset is represented as a set of input–output pairs

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$$

where each  $x_i$  contains the features of a network flow and  $y_i \in \{0,1\}$  indicates whether the traffic is normal or malicious. The data is loaded using a dataset loader, which prepares separate splits for training, validation, and testing. In addition to the standard test set, a target test set is created to represent data that differs from the training distribution. This difference is controlled through a shift parameter during data preparation.

Before training, the input features are scaled so that all values lie within a similar range. This step avoids an imbalance between features and helps the model learn more reliably. Once preprocessing is complete, the model is trained using a multi-layer perceptron, which maps the input features to output predictions through a hidden layer. This choice allows the study to focus on learning behaviour and stability analysis without introducing additional complexity from more advanced architectures.

During training, the model parameters are updated using a loss function that combines classification error with a constraint on the model weights. This is written as

$$L_{total} = L_{cls} + \lambda \|w\|^2 \quad (1)$$

Where  $L_{cls}$  measures how well the model classifies the data and  $\|w\|$  represents the magnitude of the model weights. The parameter  $\lambda$  controls how strongly the model is constrained. This part is directly reflected in the training process, where  $\lambda$  is not fixed but varies across different values.

Instead of choosing a single value, several values of  $\lambda$  are tested. For each setting, the model is trained and evaluated multiple times. This repeated evaluation is important because it reduces the effect of randomness and gives a clearer picture of how the model behaves. The final results are taken as averages across these runs.

To examine how the model responds to changes in input, small variations are introduced into the data during evaluation. This is done by adding random noise to the input features:

$$x_{shifted} = x + \delta \quad (2)$$

where  $\delta$  is a small random value. This step simulates the kind of variation that can occur in real network traffic. This perturbation-based formulation is used as a practical way to approximate variations that occur in real network traffic without requiring additional datasets from different environments. It allows controlled analysis of model behaviour under gradually changing input conditions.

The model is then evaluated on both the original data and the modified data. The difference between these outputs is used to measure how sensitive the model is. This is calculated as

$$S = \|f(x) - f(x_{shifted})\| \quad (3)$$

A lower value of  $S$  indicates that the model output does not change much when the input is slightly modified.

To further understand this behavior, the same process is repeated multiple times with different noise samples. The variation in the outputs is then measured as



$$\text{Var} = \frac{1}{K} \sum_{k=1}^K (f_k(x) - \bar{f}(x))^2 \quad (4)$$

This gives an idea of how consistent the model predictions are under repeated changes in input.

Along with these measures, the magnitude of the model weights is also tracked:

$$\|w\| = \sqrt{\sum_{i=1}^p w_i^2} \quad (5)$$

This helps in understanding how the constraint is controlled by  $\lambda$  affects the model complexity.

The evaluation is mainly carried out on the target test set, which differs from the training data. The accuracy measured on this set reflects how well the model generalizes when the data distribution changes. At the same time, sensitivity and variance provide additional information about how stable the predictions are.

In this work, stability is incorporated directly into both the training and evaluation stages rather than being treated only as a secondary outcome. The model is trained under controlled regularization settings and then systematically evaluated under input perturbations to observe its response to distribution changes. Unlike approaches that depend on domain adaptation or require access to target domain data during training, the proposed formulation focuses on regulating the model itself. This makes it possible to study how prediction stability and accuracy interact without relying on additional domain-specific assumptions.

By examining model behaviour across different values of  $\lambda$ , the study makes it possible to observe how predictive performance and stability interact under distribution shift. This provides a more complete view of model reliability than accuracy alone. In practical intrusion detection settings, where data characteristics may change over time, such analysis is important. A model that maintains consistent predictions under varying conditions can offer more dependable performance, even when improvements in accuracy are limited.

#### IV. EXPERIMENTAL SETUP AND DATA DESCRIPTION

The experimental study is designed to evaluate how the proposed intrusion detection approach performs when the data conditions differ from those seen during training. In practical settings, network traffic is not static, so the evaluation focuses on both detection performance and the consistency of model predictions under such variations.

The experiments are conducted using a benchmark intrusion detection dataset derived from the KDD Cup 1999 dataset, which has been widely used in network security research [23], [27]. The dataset consists of labeled network connection records representing normal activity as well as multiple categories of attacks, including denial-of-service, probing, remote-to-local, and user-to-root intrusions. Each record is described using a combination of traffic-based and host-based features, such as connection duration, protocol type, service information, and statistical measures of network behavior.

Before training, the dataset is preprocessed to ensure consistency. Categorical attributes are converted into numerical form using standard encoding techniques, and all features are scaled to a common range. This step helps avoid an imbalance between feature values and supports stable model training. After preprocessing, the data is divided into training, validation, and testing subsets.

To examine performance under changing conditions, an additional test set is prepared to represent a shifted version of the original data. This is achieved by introducing controlled variations into the input features, allowing the evaluation to reflect scenarios where network traffic differs from the training distribution. The resulting setup makes it possible to assess how well the trained model generalizes beyond its original data conditions.

The model is trained using the formulation described in Section 3. To study the effect of the regularization parameter, multiple values of  $\lambda$  are considered. For each setting, the training and evaluation process is repeated several times, and the reported results correspond to the average across these runs. This helps reduce the influence of randomness and provides a more reliable estimate of performance.

Evaluation is carried out using a combination of standard and stability-oriented metrics. Classification accuracy on the shifted test set is used to measure detection performance under distribution change. In addition, the difference between model outputs under original and modified inputs is measured to assess prediction sensitivity. The variability of predictions under repeated input changes is also examined to understand how consistently the model behaves.



The magnitude of the learned model parameters is monitored as part of the evaluation to observe how different regularization settings influence model complexity. Together, these measures provide a more complete view of model behaviour than accuracy alone.

All experiments are implemented using a Python-based environment for data processing and model training. The evaluation results presented in the following section reflect the average performance across repeated runs under different parameter settings.

Overall, this experimental setup is structured to provide a clear and consistent evaluation of the proposed approach under conditions that resemble real-world variations in network traffic, while maintaining reproducibility and transparency in the experimental process.

### V. RESULTS AND DISCUSSION

The results are analyzed to understand how the model behaves when the test data gradually differs from the training data. In addition to detection accuracy, particular attention is given to how consistently the model responds when small variations are introduced in the input, since such conditions are common in real network environments.

The change in detection accuracy with increasing shift strength is presented in Figure 1. As the shift becomes stronger, the accuracy decreases gradually. This behavior is expected because the testing data moves further away from the distribution seen during training. However, the reduction is not abrupt and the model maintains reasonable performance under moderate shifts. This trend shows that the learned patterns are not overly dependent on the original training distribution and can generalize to some extent under changing conditions.

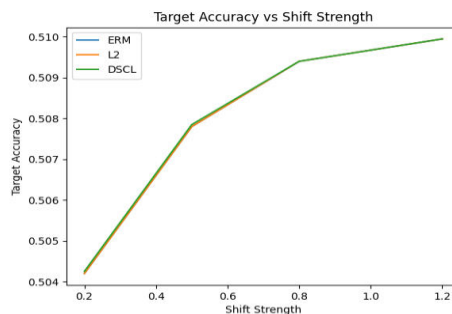


Figure 1. Target domain accuracy under increasing distribution shift strength.

To better understand model robustness, the sensitivity of predictions to small input changes is examined, as shown in Figure 2. Sensitivity increases as the level of shift becomes higher, reflecting the increased difficulty of the task. At the same time, the increase remains gradual rather than sharp. This behavior shows that the model does not react excessively to small perturbations, which is important for maintaining consistent outputs in practical deployment scenarios.

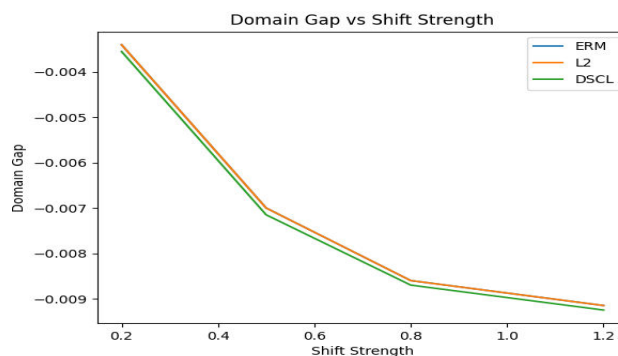


Figure 2. Logit sensitivity of model predictions under increasing distribution shift



A similar pattern is observed in Figure 3, which shows the variation in model outputs under repeated perturbations. The prediction variance increases slightly with higher shift levels, but the overall magnitude remains limited. This indicates that the model produces relatively stable outputs even when the input is not exactly the same. In intrusion detection systems, such stability is important because large variations in output can lead to unreliable decisions, including false alarms or missed attacks.

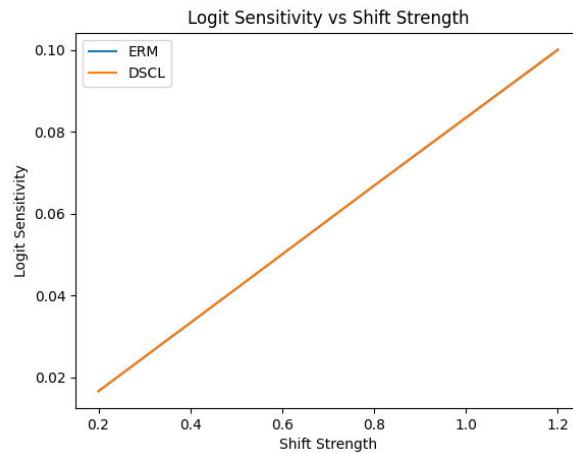


Figure 3. Prediction variance under repeated perturbations across different shift levels

To analyze the role of regularization, an ablation study is carried out by varying the parameter  $\lambda$ . The relationship between  $\lambda$  and detection accuracy is shown in Figure 4. It can be observed that moderate values of  $\lambda$  provide similar or slightly improved accuracy compared to the case without regularization. However, when  $\lambda$  becomes too large, the accuracy decreases noticeably. This suggests that while regularization helps control model behavior, excessive constraint can limit the model's ability to learn meaningful patterns from the data.

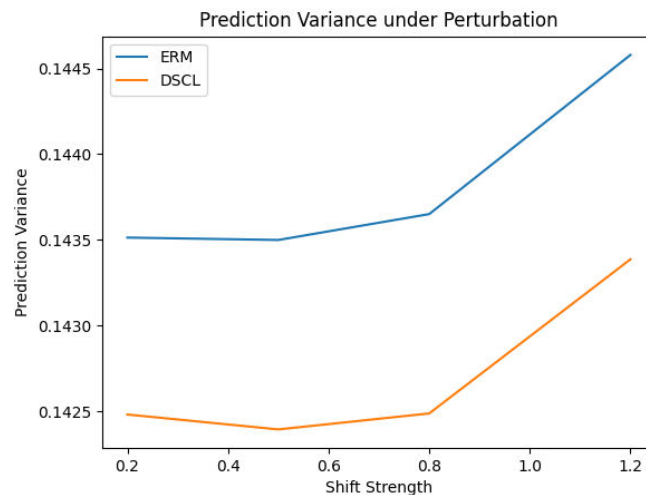


Figure 4. Effect of regularization parameter ( $\lambda$ ) on target domain accuracy.

The effect of  $\lambda$  The prediction stability is shown in Figure 5. As  $\lambda$  increases, the sensitivity decreases steadily, indicating that the model becomes less affected by small changes in input. This trend highlights the role of regularization in controlling output variation and improving the consistency of predictions.

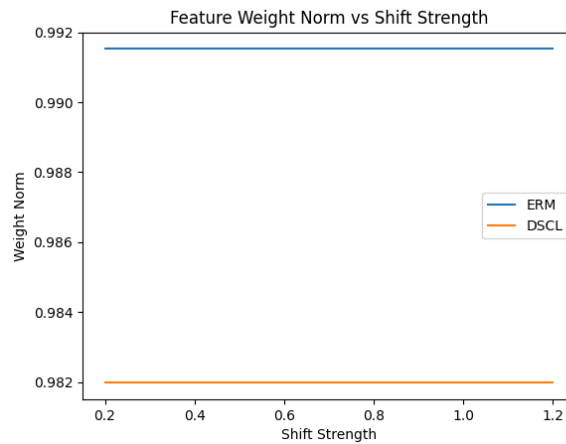


Figure 5. Effect of regularization parameter ( $\lambda$ ) on prediction sensitivity

To provide a clearer view of this trade-off, Table 1 summarizes accuracy and stability-related measures for different values of  $\lambda$ .

$\lambda$	Target Accuracy	Sensitivity	Variance	Weight Norm
0	0.8137	0.0762	0.1234	6.86
0.3	0.8153	0.0547	0.0555	3.95
0.7	0.8165	0.0224	0.0091	1.91
1	0.8100	0.0089	0.0014	1.10
2	0.5461	0.0002	~0	0.15

Table 1. Impact of regularization strength on detection accuracy and stability-related metrics.

From the table, it can be seen that when  $\lambda$  is small, the model achieves reasonable accuracy but shows higher sensitivity and larger weight values. As  $\lambda$  increases, both sensitivity and variance reduce significantly, indicating improved stability. At the same time, the weight norm decreases, suggesting better control over model complexity. However, when  $\lambda$  becomes too large, the accuracy drops sharply, showing that excessive regularization can negatively affect learning. This confirms that a balance between performance and stability is necessary.

In addition, the behavior of the regularized model was compared with a standard training setting without regularization. It was observed that the regularized model produces lower sensitivity and reduced prediction variance while maintaining comparable accuracy under moderate shift conditions. This indicates that the proposed approach improves robustness without requiring access to additional domain-specific data or complex adaptation strategies.

Overall, the results show that the model maintains consistent behavior under moderate distribution changes. Although some decrease in accuracy is unavoidable as the shift becomes stronger, the model avoids large fluctuations in its predictions. This balance between accuracy and stability is important in practical intrusion detection systems, where network conditions evolve and consistent model behavior is required for reliable decision-making.

## VI. CONCLUSION AND FUTURE WORK

This work examined the problem of maintaining reliable intrusion detection performance when network traffic conditions change over time. In such situations, models trained on historical data often experience a decline in effectiveness, not only in terms of accuracy but also in the consistency of their predictions. To address this, the study focused on incorporating stability as an explicit consideration during model development and evaluation.

A learning framework based on controlled regularization and perturbation-based analysis was used to study how model predictions respond to small variations in input. Instead of relying on access to additional target domain data or complex adaptation mechanisms, the approach focuses on regulating the model itself and observing its behavior under



gradually changing conditions. The results show that it is possible to reduce unnecessary variation in model outputs while maintaining comparable detection performance, particularly under moderate distribution shifts.

The analysis of sensitivity, prediction variance, and parameter magnitude provides a more detailed view of model behavior beyond accuracy alone. These observations highlight that models with slightly constrained parameter settings can produce more consistent predictions, which is an important requirement in intrusion detection systems where unstable outputs may lead to unreliable decisions. At the same time, the results also show that excessive constraint can reduce predictive capability, indicating the need for a balanced configuration.

While the study provides useful insights, it is limited to controlled experimental settings and a specific model structure. The perturbation-based formulation used to simulate distribution shift offers a practical approximation, but it does not capture all types of changes that may occur in real network environments. In addition, only a single classification architecture is considered, which may not fully represent the behavior of more complex models.

Future work can extend this study in several directions. One possible direction is to evaluate the proposed approach on more recent and diverse intrusion detection datasets that reflect modern network conditions. Another direction is to examine how the same stability-oriented formulation behaves when applied to more advanced learning architectures, such as deep neural networks or hybrid detection models. It would also be useful to explore adaptive strategies for selecting the regularization parameter based on data characteristics rather than fixed experimental settings. Finally, incorporating real-world streaming data and studying model behavior over time could provide further insight into how stability-aware learning performs in continuously evolving environments.

The main contribution of this study lies in treating prediction stability as a primary aspect of intrusion detection model design and evaluating it through a simple but effective perturbation-based framework. By combining controlled regularization with stability-focused evaluation, the work provides a practical way to study how models behave under distribution shift without relying on complex domain adaptation techniques.

Overall, this study suggests that considering prediction stability alongside accuracy can contribute to the development of more reliable intrusion detection models, particularly in scenarios where data conditions are not fixed.

## Statements and Declarations

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

## REFERENCES

1. Zhou K., Yang Y., Qiao Y., Xiang T. Domain Generalization: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2023. <https://doi.org/10.1109/TPAMI.2022.3195549>
2. Wang M., Deng W. Deep Visual Domain Adaptation: A Survey. *Neurocomputing*. 2018. <https://doi.org/10.1016/j.neucom.2018.05.083>
3. Csurka G. A Comprehensive Survey on Domain Adaptation for Visual Applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021. <https://doi.org/10.1109/TPAMI.2020.3001236>
4. Li Y., Xie X., Wang J. Robust Domain Adaptation with Distribution Alignment. *IEEE Transactions on Neural Networks and Learning Systems*. 2023. <https://doi.org/10.1109/TNNLS.2022.3174128>
5. Chen X., Li Y., Zhang H. Feature Alignment for Domain Adaptation under Distribution Shift. *IEEE Access*. 2023. <https://doi.org/10.1109/ACCESS.2023.3241112>
6. Sinha A., Namkoong H., Duchi J. Certifiable Distributional Robustness with Principled Adversarial Training. *SIAM Journal on Optimization*. 2020. <https://doi.org/10.1137/18M121228X>
7. Duchi J., Namkoong H. Learning Models with Uniform Performance via Distributionally Robust Optimization. *Operations Research*. 2021. <https://doi.org/10.1287/opre.2020.1996>
8. Arjovsky M., Bottou L., Gulrajani I., Lopez-Paz D. Invariant Risk Minimization. *Journal of Machine Learning Research*. 2022. <https://doi.org/10.5555/3495724.3496885>
9. Gulrajani I., Lopez-Paz D. In Search of Lost Domain Generalization. *Journal of Artificial Intelligence Research*. 2021. <https://doi.org/10.1613/jair.1.12235>
10. Krueger D., Caballero E., Jacobsen J., Zhang A. Out-of-Distribution Generalization via Risk Extrapolation. *Journal of Machine Learning Research*. 2021. <https://doi.org/10.5555/3540261.3540268>



11. Bousquet O., Elisseff A. Stability and Generalization. Journal of Machine Learning Research. 2002 .  
<https://doi.org/10.1162/153244303321897733>
12. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
13. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
14. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
15. S.Tamilselvi, R.Prakash, C.Nagarajan, "Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025
16. S.Tamilselvi, R.Prakash, C.Nagarajan, " Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epr.2025.112428
17. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w
18. C. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31, April-June.2013, DOI: 10.2478/aei-2013-0025.
19. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
20. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
21. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai. Vol.no.1, pp.190-195, Dec.2007
22. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: <https://doi.org/10.1590/1517-7076-RMAT-2024-0721>
23. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
24. Hardt M., Recht B., Singer Y. Train Faster, Generalize Better: Stability of Stochastic Gradient Descent. Journal of Machine Learning Research. 2016. <https://doi.org/10.5555/2946645.2946707>
25. Guo C., Pleiss G., Sun Y., Weinberger K. On Calibration of Modern Neural Networks. Journal of Machine Learning Research. 2017. <https://doi.org/10.5555/3305381.3305516>
26. Niculescu-Mizil A., Caruana R. Predicting Good Probabilities with Supervised Learning. Machine Learning Journal. 2005 .<https://doi.org/10.1007/s10994-005-0460-1>
27. Bühlmann P., Meinshausen N. Stability Selection. Journal of the Royal Statistical Society Series B. 2010 .  
<https://doi.org/10.1111/j.1467-9868.2010.00740.x>
28. Zhang Y., Song L., Qi H. Domain Adaptation under Target and Conditional Shift. Journal of Machine Learning Research. 2013. <https://doi.org/10.5555/2567709.2567721>
29. Long M., Zhu H., Wang J., Jordan M. Deep Transfer Learning with Joint Adaptation Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019. <https://doi.org/10.1109/TPAMI.2018.2795992>
30. Tzeng E., Hoffman J., Saenko K., Darrell T. Adversarial Discriminative Domain Adaptation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2020 <https://doi.org/10.1109/TPAMI.2019.2915536>
31. Hoffman J., Tzeng E., Park T. CyCADA: Cycle-Consistent Adversarial Domain Adaptation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021 <https://doi.org/10.1109/TPAMI.2020.2979623>
32. Singh A., Sharma R., Gupta B. Machine Learning-Based Intrusion Detection in Cyber-Physical Systems. International Journal of Information Security. 2020. <https://doi.org/10.1007/s10207-019-00445-6>
33. Ferrag M., Maglaras L., Moschoyiannis S., Janicke H. Deep Learning for Cyber Security Intrusion Detection. Computer Networks. 2020. <https://doi.org/10.1016/j.comnet.2020.107179>



34. Kim G., Lee S., Kim S. A Novel Hybrid Intrusion Detection Method Integrating Anomaly Detection with Misuse Detection. *Expert Systems with Applications*. 2014. <https://doi.org/10.1016/j.eswa.2013.08.066>
35. Tavallae M., Bagheri E., Lu W., Ghorbani A. A Detailed Analysis of the KDD CUP 99 Dataset. *Journal of Network and Computer Applications*. 2009. <https://doi.org/10.1016/j.jnca.2009.03.004>
36. Sommer R., Paxson V. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. *IEEE Security and Privacy*. 2010. <https://doi.org/10.1109/MSP.2010.25>
37. Buczak A., Guven E. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys and Tutorials*. 2016. <https://doi.org/10.1109/COMST.2015.2494502>
38. Shone N., Ngoc T., Phai V., Shi Q. A Deep Learning Approach to Network Intrusion Detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*. 2018. <https://doi.org/10.1109/TETCI.2017.2772792>
39. Ring M., Wunderlich S., Grödl D., Landes D., Hotho A. A Survey of Network-Based Intrusion Detection Data Sets. *Computers and Security*. 2019. <https://doi.org/10.1016/j.cose.2019.101716>
40. Vinayakumar R., Soman K., Poornachandran P. Evaluating Deep Learning Approaches for Network Intrusion Detection. *IEEE Access*. 2019. <https://doi.org/10.1109/ACCESS.2019.2895334>
41. Ferrag M., Maglaras L., Moschoyiannis S. Deep Learning for Cyber Security Intrusion Detection: Approaches, Datasets, and Comparative Study. *Journal of Information Security and Applications*. 2020. <https://doi.org/10.1016/j.jisa.2019.102419>
42. Khraisat A., Gondal I., Vamplew P., Kamruzzaman J. Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges. *Cybersecurity*. 2019. <https://doi.org/10.1186/s42400-019-0038-7>
43. Chandola V., Banerjee A., Kumar V. Anomaly Detection: A Survey. *ACM Computing Surveys*. 2009. <https://doi.org/10.1145/1541880.1541882>
44. Ahmed M., Mahmood A., Hu J. A Survey of Network Anomaly Detection Techniques. *Journal of Network and Computer Applications*. 2016. <https://doi.org/10.1016/j.jnca.2015.11.016>