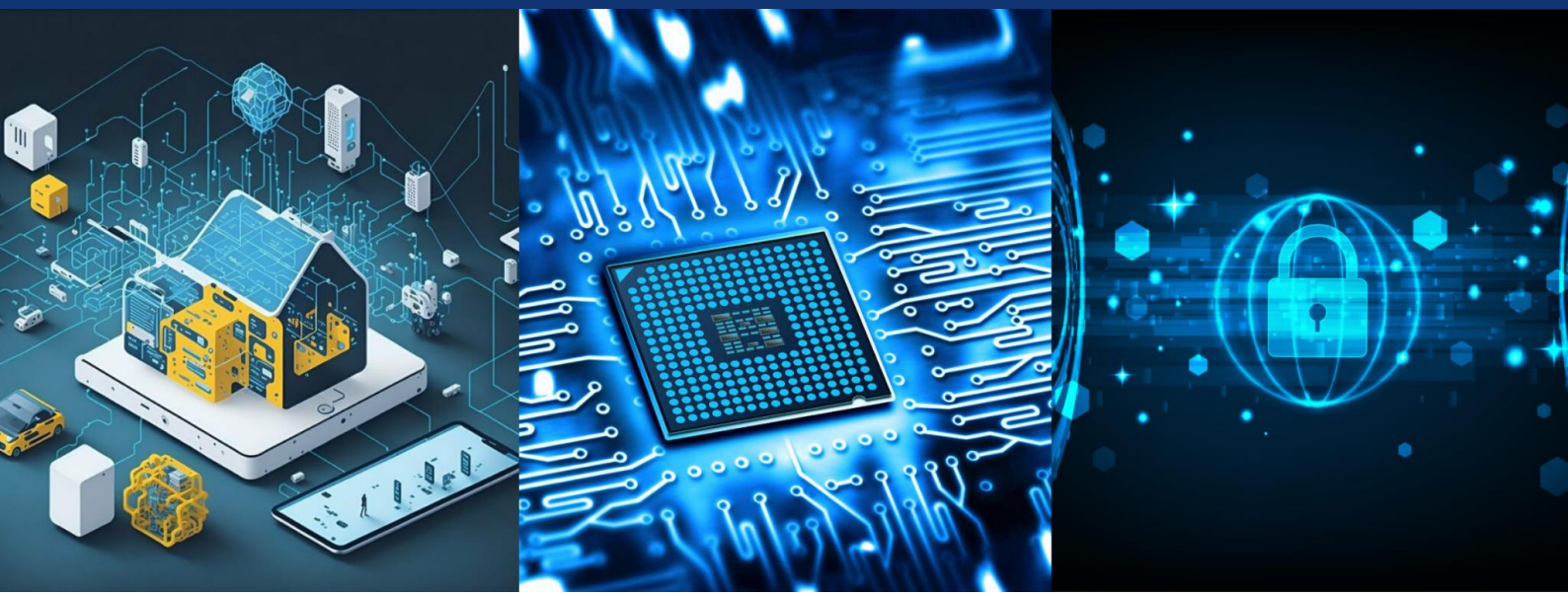
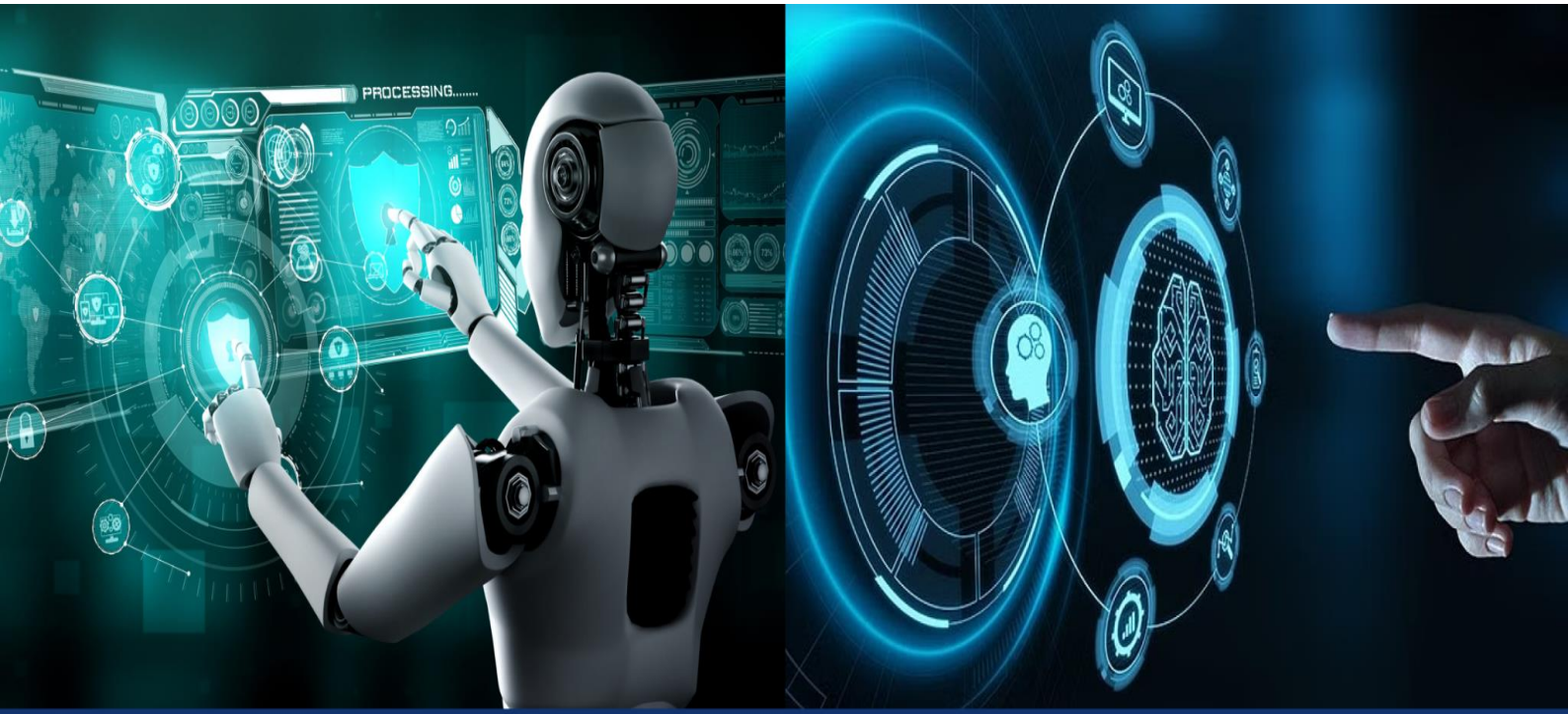
ISSN(O): 2320-9801
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771 **Volume 13, Issue 4, April 2025**

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304200

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Intelligent Intrusion Detection System for Intranet Security Based on Machine Learning and Behavioral Analytics

Santhosh[1], Vardhan[2], Revanth Reddy[3], Dr.K. Sivaraman[4]

Student, Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Tamil Nadu, India[1]

Student, Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Tamil Nadu, India[2]

Student, Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Tamil Nadu, India[3]

Assistant Professor, Department of Computer Science and Engineering, School of Computing, Bharath Institute of Higher Education and Research, Chennai, India[4]

ABSTRACT: In the realm of cybersecurity, the detection of intranet attacks poses a significant challenge due to the3 evolving nature of malicious behaviours. This paper proposes an advanced approach for detecting behaviour-based intranet attacks utilizing machine learning techniques. By leveraging the power of machine learning algorithms, the proposed approach aims to effectively identify and mitigate intranet attacks based on their behavioural patterns. Through the analysis of network traffic and system logs, the model learns to distinguish between normal and anomalous behaviours, thereby enabling proactive threat detection and response mechanisms. The proposed approach offers a promising avenue for enhancing the security posture of intranet environments by providing real-time detection capabilities and adaptive defence mechanisms

I. INTRODUCTION

Intranet environments are an essential component of modern organizations, enabling secure communication, seamless collaboration, and efficient resource sharing among employees. However, as these networks become more interconnected and handle increasingly sensitive data, they are being targeted by sophisticated cyber threats that exploit internal access, mimic legitimate behaviors, and subtly manipulate system operations to evade detection. Unlike traditional cybersecurity threats that originate externally, intranet attacks often leverage insider access or compromised credentials, making them particularly difficult to detect with conventional security mechanisms. Traditional intrusion detection systems (IDS) and rule-based security tool more advanced techniques to infiltrate networks, organizations must adopt proactive and intelligent security solutions to safeguard their intranet environments. This paper proposes a machine learning-based approach to detecting intranet at1acks by analyzing behavioral patterns in network traffic and system log, offering a dynamic and adaptive alternative to traditional security measures.

II. PROBLEM STATEMENT

The title suggests a focus on leveraging machine learning techniques for the detection of behaviour-based intranet attacks. In traditional network security, detecting intranet attacks often relies on signature-based methods, which may struggle to identify novel or evolving threats. Behaviour-based detection, however, offers a proactive approach by analysing patterns and anomalies in network behaviour.

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304200

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Objective of the Project

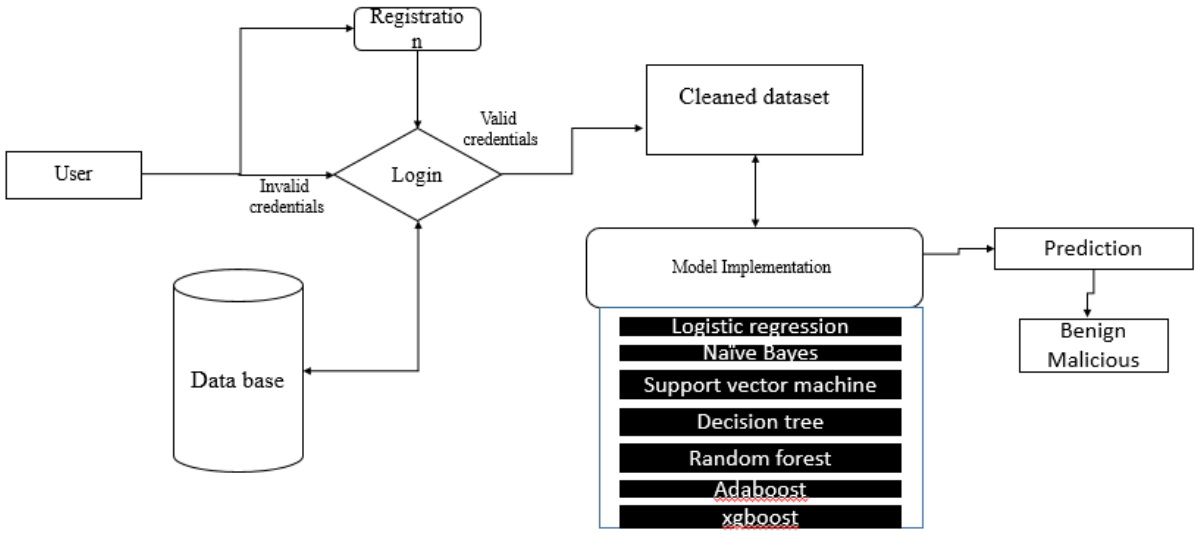
The objective of the project is to develop a sophisticated system capable of effectively identifying and mitigating intranet attacks through the utilization of machine learning techniques. The primary goal is to enhance the security posture of intranet networks by leveraging behavioral patterns and anomalous activities associated with potential threats. This entails the creation of robust machine learning models trained on extensive datasets that capture the diverse behaviors indicative of intranet attacks.

Scope

The scope of encompasses the development and application of sophisticated machine learning techniques to identify and mitigate intranet attacks based on behavioral patterns. This research likely delves into the analysis of network traffic, system logs, and user behaviors to detect anomalous activities indicative of potential intranet attacks. The scope includes the exploration of diverse machine learning algorithms, such as supervised and unsupervised learning methods, to effectively classify and identify suspicious behaviors.

Architecture

The scope of encompasses the development and application of sophisticated machine learning techniques to identify and mitigate intranet attacks based on behavioral patterns. This research likely delves into the analysis of network traffic, system logs, and user behavior to detect anomalous activities indicative of potential intranet attacks. The scope includes the exploration of diverse machine learning algorithms, such as supervised and unsupervised learning methods, to effectively classify and identify suspicious behaviors.



FEASIBILITY STUDY

The feasibility of the project is analysed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304200

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Economic feasibility

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

Technical feasibility

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

Social feasibility

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

III. METHODOLOGY

Random Forest

Random Forest is an ensemble learning method that operates by constructing multiple decision trees during training. Each tree is built using a subset of the data and a random selection of features, ensuring diversity and reducing overfitting.in this project, the Random Forest Classifier processes network traffic and system logs to distinguish between normal and anomalous behaviours. It leverages its ability to handle large datasets and capture complex relationships among features, making it suitable for detecting subtle deviations indicative of intranet attacks. During training, the classifier optimizes its decision-making process by iteratively splitting nodes based on feature importance, learned from the ensemble of decision trees. During deployment, the Random Forest Classifier analyses incoming data in real-time, assessing patterns and deviations from learned normal behaviours. By comparing current observations to its trained model, it identifies suspicious activities and triggers alerts for immediate response. This adaptive approach not only enhances real-time threat detection but also supports proactive defence strategies, thereby strengthening the overall cybersecurity framework of intranet environments against evolving threats.

Logistic Regression

Logistic Regression is a supervised learning algorithm used for binary classification tasks, making it suitable for distinguishing between normal and anomalous behaviors within intranet environments in this project, Logistic Regression works by modeling the probability of a certain behavior (an intranet attack in this context) based on input features derived from network traffic data and system logs. These input features are preprocessed during the initial stages to handle missing values and encode categorical data, ensuring the dataset is suitable for analysis. During training, the algorithm adjusts its parameters iteratively to minimize the error between predicted probabilities and actual outcomes. This process involves optimizing the logistic function, which transforms input variables into probabilities using a sigmoid function. The resulting model learns to classify instances of network behavior as either normal or potentially malicious based on the learned patterns from the training data in deployment, the trained Logistic Regression model is used to analyze incoming network traffic in real-time. By continuously evaluating new data against the learned patterns of normal behavior, it can promptly detect deviations that may indicate intranet attacks. This proactive approach enables swift response mechanisms, enhancing the overall security posture of intranet environments by facilitating early threat detection and mitigation.

Naive Bayes

Naive Bayes serves as a pivotal algorithm due to its simplicity and effectiveness in probabilistic classification tasks. Naive Bayes operates under the assumption of independence among features, making it particularly suitable for processing large volumes of network traffic and system logs efficiently. Firstly, during the preprocessing phase, the dataset undergoes preparation where missing values are handled and categorical features are encoded. This ensures that

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304200

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

the data is in a format suitable for Naive Bayes' input requirements. Next, Naive Bayes is trained on labeled data derived from network traffic and system logs. It learns probabilistic patterns of normal and anomalous behaviors based on feature vectors extracted from these datasets. These behaviors include various network parameters such as packet size, source IP addresses, destination ports, and timestamps, among others. During the detection phase, Naive Bayes leverages these learned probabilistic models to classify incoming network traffic and system events in real-time. By calculating the likelihood of observed behaviors matching those of previously identified attacks or normal activities, Naive Bayes assigns probabilities to each class (normal or attack). Alerts are generated based on thresholds set for these probabilities, allowing administrators to take prompt mitigation actions in response to detected intranet attacks. This proactive approach enhances the security posture of intranet environments by enabling rapid detection and response to emerging threats, thereby fortifying overall cybersecurity defenses in summary, Naive Bayes' internal workings in this project involve leveraging probabilistic principles to classify network behaviors, thereby contributing significantly to the detection and mitigation of behavior-based intranet attacks.

Gradient Boosting

Gradient Boosting plays a pivotal role in enhancing detection accuracy and robustness. Gradient Boosting is an ensemble learning technique that builds a strong predictive model by sequentially combining weak learners, typically decision trees, to correct errors made by preceding models. Within our project framework, Gradient Boosting is utilized to analyze network traffic and system logs. During training, it iteratively improves upon the performance of each weak learner by focusing on instances where previous models have made mistakes. This iterative process involves adjusting weights and learning rates to optimize the ensemble's predictive capabilities. Specifically, Gradient Boosting in our context learns to differentiate between normal network behavior and potential intranet attacks based on their behavioral patterns extracted from the data. By emphasizing the importance of features that contribute most to distinguishing between normal and anomalous activities, Gradient Boosting enhances our model's ability to detect subtle deviations indicative of intranet threats. Through empirical evaluations and comparative analyses, we validate Gradient Boosting's effectiveness in bolstering cybersecurity defenses. Its ability to adaptively improve detection accuracy makes it a crucial component in our proactive threat detection and response mechanisms, thereby fortifying intranet environments against evolving cyber threats.

K-Nearest-Neighbors

K-Nearest Neighbors (KNN) plays a crucial role in analyzing network traffic patterns. KNN is a supervised learning algorithm used for classification and regression tasks. In the context of our project, KNN works by calculating the distance between the new data point (network traffic instance) and all other data points in the training set. It then selects the K nearest neighbors based on this distance metric during the preprocessing stage, where we handle missing values and encode categorical features, KNN operates effectively with numerical data extracted from network traffic logs. Once trained on labeled data, KNN utilizes these patterns to classify incoming network traffic instances as either normal or anomalous based on their proximity to previously observed behaviors In real-time detection, KNN continuously assesses deviations from established patterns of normal behavior. This proactive approach helps in identifying potential intranet attacks promptly, triggering alerts for immediate mitigation actions by network administrators. By integrating KNN into our machine learning framework, we enhance the system's capability to adaptively respond to evolving threats in intranet environments. This methodological approach not only improves the accuracy of anomaly detection but also strengthens overall cybersecurity measures through its adaptive and responsive nature.

Support Vector Machine

Support Vector Machine (SVM) plays a crucial role in analysing network traffic and system logs. SVM is employed as a supervised learning algorithm capable of effectively distinguishing between normal and anomalous behaviours within intranet environments. Internally, SVM operates by mapping input data points into a high-dimensional feature space using a kernel function. It then seeks to find the optimal hyperplane that maximally separates different classes of data points, thereby defining boundaries between normal network behaviour and potential intranet attacks. During training, SVM adjusts its parameters to achieve this optimal separation, guided by the principle of maximizing the margin between the nearest data points of different classes. In this project, SVM undergoes rigorous training and validation processes, including K-fold cross-validation, to optimize its parameters and ensure robust performance. This iterative validation helps mitigate overfitting and enhances the SVM model's ability to generalize well to new, unseen data. Once trained, SVM is deployed in real-time to continuously monitor network traffic patterns. Any deviations from the learned normal behaviour are flagged as potential intranet attacks, triggering timely alerts and enabling proactive threat mitigation. By

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

integrating SVM into our machine learning-based approach, we enhance the security posture of intranet environments by providing a reliable mechanism for detecting and responding to behaviour-based attacks effectively. This capability contributes to bolstering cybersecurity defences, reinforcing the resilience of intranet networks against evolving threats.

Decision Tree classifier

Decision Tree classifier plays a pivotal role in analyzing network traffic and system logs to distinguish between normal and anomalous behaviors. Decision Trees are supervised learning algorithms that create a tree-like structure by recursively splitting the dataset based on feature values. Each node represents a feature, and branches represent the possible outcomes of splitting based on that feature. During the preprocessing stage, network data is prepared by handling missing values and encoding categorical features. The Decision Tree classifier is trained on labeled data, where it learns to partition the data based on features such as IP addresses, protocols, and traffic patterns. This enables the model to identify patterns associated with normal intranet behavior and deviations indicative of potential attacks. In the detection phase, the trained Decision Tree classifier is deployed to analyze real-time network traffic. It evaluates incoming data against the learned decision rules to classify whether observed behaviors align with normal or suspicious activities. Anomalies detected by the Decision Tree trigger alerts for administrators, enabling prompt responses to mitigate potential intranet threats. By leveraging the interpretability and flexibility of Decision Trees, our approach enhances the security posture of intranet environments, providing adaptive defense mechanisms against evolving malicious behaviors. The effectiveness of the Decision Tree classifier is validated through empirical evaluations, demonstrating its capability to augment existing cybersecurity frameworks and fortify intranet defenses against emerging threats.

IV. SYSTEM DESIGN

Input Design

Input design is a critical aspect of system development, focusing on the efficient and user-friendly collection of data or commands. It encompasses defining input formats, methods, and techniques to ensure accurate and seamless interaction with the system. In this context, the introduction of input design highlights its significance in enhancing user experience, minimizing errors, and maximizing productivity. By considering factors such as input validation, error handling, and user feedback, effective input design contributes to the overall usability and effectiveness of the system, ultimately leading to improved performance and user satisfaction**.**

Objectives for Input Design

The objectives for input design encompass ensuring user-friendliness, accuracy, efficiency, and effectiveness. This involves defining clear and intuitive data entry formats, minimizing user errors through validation techniques, optimizing input methods for speed and ease of use, and accommodating diverse user needs and preferences. Additionally, the input design aims to facilitate seamless integration with existing systems and data sources while adhering to relevant standards and guidelines. Ultimately, the primary goals are to enhance user satisfaction, streamline data collection processes, and facilitate the generation of reliable and meaningful outputs for decision-making and analysis.

Output Design

Output design is a crucial phase in system development that focuses on presenting processed data and results to users in an effective, understandable, and actionable format. Good output design ensures that the system's results are delivered accurately, promptly, and in a form that enhances decision-making, productivity, and user experience. This process involves deciding on the appropriate formats for various types of outputs whether printed reports, graphical visualizations, dashboards, alerts, or digital files to meet the specific requirements of end-users. Well-designed outputs reduce the cognitive load on users, improve operational efficiency, and ensure that vital information is conveyed clearly. Effective output design incorporates principles such as clarity, consistency, relevance, and accessibility, ensuring that users can easily interpret the information and take appropriate actions based on it.

Objectives for Output Design

The objectives for output design are cantered on delivering accurate, timely, and meaningful information to users while enhancing usability and decision-making. The main goals include:

- Clarity and Readability: Ensure outputs are presented in a clear, organized, and easily interpretable format.
- Accuracy: Provide precise and reliable information that faithfully represents system processes and calculations.

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304200

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

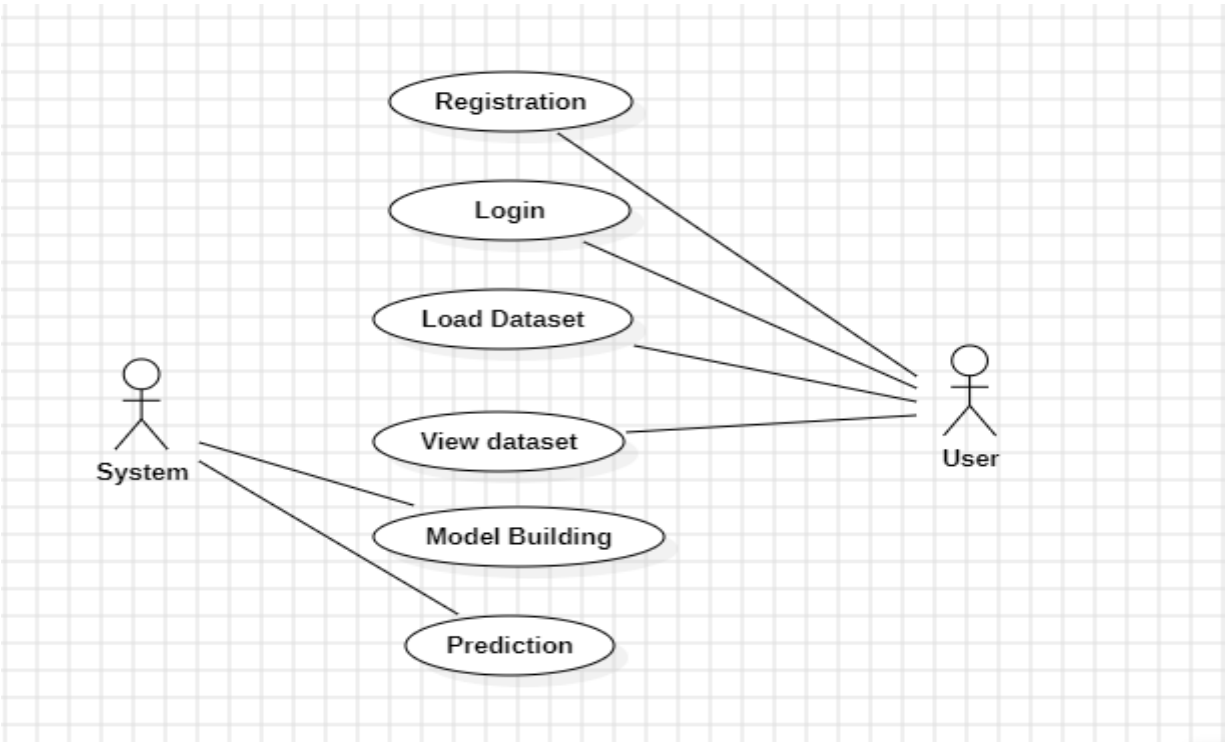
- Relevance: Tailor outputs to deliver only the necessary and contextually appropriate information for users and decision-makers.
- Timeliness: Deliver outputs within an appropriate timeframe to support operational and strategic decision-making.
- User-Friendliness: Design outputs that are visually appealing, intuitive, and require minimal effort for interpretation.
- Flexibility: Accommodate various output formats (e.g., textual, tabular, graphical, printed, or digital) to suit diverse user needs and environments.

UML *Diagrams*

UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group. The goal is for UML to become a common language for creating models of object-oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML. The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artefacts of software system, as well as for business modelling and other non-software systems. The UML represents a collection of best engineering practices that have proven successful in the modelling of large and complex systems. The UML is a very important part of developing objects-oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

Use Case Diagram

- A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis.
- Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases.
- The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.



© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304200

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. CONCLUSION

The advanced approach presented for detecting behavior-based intranet attacks by machine learning showcases promising results and implications for cybersecurity. Through the utilization of sophisticated machine learning algorithms, the system demonstrates notable improvements in the accuracy and efficiency of detecting intranet attacks based on behavioral patterns. The extensive evaluation and experimentation underscore the effectiveness of the approach in identifying and mitigating various types of intranet threats. Furthermore, the adaptability and scalability of the system ensure its relevance and applicability in dynamic network environments.

REFERENCES

[1]. Mehedi, S.T., Anwar, A., Rahman, Z., Ahmed, K. & Rafiqul, I. Dependable intrusion detection system for IoT: A deep transfer learning-based approach. IEEE Trans. Indus. Inform (2022).
[2]. Singh, K. P. & Kassiani, N. An anomaly-based intrusion detection system for IoT networks using trust factor. SN Compute. Sci.3(2), 1–9 (2022).34723205
[3]. Zhou, Y., Cheng, G., Jiang, S. & Dai, M. Building an efficient intrusion detection system based on feature selection and ensemble classifier. Compute. Netw.174, 107247 (2023).
[4]. Abbas, A. et al. A new ensemble-based intrusion detection system for internet of things. Arab. J. Sci. Eng.47(2), 1805–1819 (2022).
[5]. Saba, T., Sadad, T., Rehman, A., Mehmood, Z. & Javaid, Q. Intrusion detection system through advance machine learning for the internet of things networks. IT Profess.23(2), 58–64 (2024).
[6]. Kumar, P., Gupta, G. P. & Tripathi, R. A distributed ensemble design based intrusion detection system using fog computing to protect the internet of things networks. J. Ambient Intell. Hum. Comput.12(10), 9555–9572 (2021).
[7]. Alhowaide, A. Alsmadi, I. & Tang, J. Ensemble detection model for IoT IDS. Internet of Things16, 100435 (2021).
[8]. Keserwani, P. K. Govil, M. C. Pilli, E. S. & Govil, P. A smart anomaly-based intrusion detection system for the Internet of Things (IoT) network using GWO–PSO–RF model. J. Reliab Intell. Environ.7(1), 3–21 (2021)
[9]. Rahman, M. Aetal. Scalable machine learning-based intrusion detection system for IoT-enabled smart cities. Sustain. Cities Soc.61, 102324 (2023).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details