Network Intrusion Detection Using AI/ML

Divyanshu Mittal
(MSA23021)

Under the supervision of

Dr. Abhinesh Kaushik

Indian Institute of Information Technology, Lucknow

November 21, 2024

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Overview

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Introduction

- Wireless Sensors Network is a network of tiny nodes to observe the environment.
- Any type of unwanted or unusual activity is called an intrusion.
- These intrusions can steal the sensitive information.
- Intrusions are the significant threat to security of the network and the sensitive data of the users.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Introduction

- Wireless Sensors Network is a network of tiny nodes to observe the environment.
- Any type of unwanted or unusual activity is called an intrusion.
- These intrusions can steal the sensitive information.
- Intrusions are the significant threat to security of the network and the sensitive data of the users.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Introduction

- Wireless Sensors Network is a network of tiny nodes to observe the environment.
- Any type of unwanted or unusual activity is called an intrusion.
- These intrusions can steal the sensitive information.
- Intrusions are the significant threat to security of the network and the sensitive data of the users.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Introduction

- Wireless Sensors Network is a network of tiny nodes to observe the environment.
- Any type of unwanted or unusual activity is called an intrusion.
- These intrusions can steal the sensitive information.
- Intrusions are the significant threat to security of the network and the sensitive data of the users.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Problem Statement

- Cybersecurity threats are increasing day-by-day, creating challenges to the traditional intrusion detection techniques.
- Machine learning models like Random Forest, Decision Trees, k-Nearest Neighbors can be used to enhance the detection accuracy.
- This project aims to build a model which can detect the intrusions efficiently.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Problem Statement

- Cybersecurity threats are increasing day-by-day, creating challenges to the traditional intrusion detection techniques.
- Machine learning models like Random Forest, Decision Trees, k-Nearest Neighbors can be used to enhance the detection accuracy.
- This project aims to build a model which can detect the intrusions efficiently.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Problem Statement

- Cybersecurity threats are increasing day-by-day, creating challenges to the traditional intrusion detection techniques.
- Machine learning models like Random Forest, Decision Trees, k-Nearest Neighbors can be used to enhance the detection accuracy.
- This project aims to build a model which can detect the intrusions efficiently.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Literature Review

Algorithm	Year	Reference	Result	Limitation
kNN	2014	W. Li, P. Yi, Y. Wu, L. Pan, J. Li	98.5%	Sensitive to the value of 'k'
SVM, kNN	2017	M. M. U. Chowdhury, F. Hammond, G. Konowicz, C. Xin, H. Wu, and J. Li	94.62%	Computational requirements are very high
Deep Learning	2019	R. Vinay, M. Alazab, K. P. Soman, P. Poorna, A. Al-Nemrat	88.7%	Requires high computational power and large dataset
kNN and sine-cosine algorithm	2021	J.-S. Pan, F. Fan, S.-C. Chu, H.-Q. Zhao, and G.-Y. Liu	99.34%	Not good for higher dimensional data
SVM and Firefly algorithm	2024	M. Karthikeyan, D. Manimegalai, and K. RajaGopal	99.34%	Not scalable, not useful for every intrusion detection

Table 1: Summary of Literature on intrusion detection

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Methodology

- Reviewed literature on intrusion detection.
- Loaded and explored the NSL-KDD dataset.
- Data preprocessing, feature exploration and feature extraction using PCA for model development.
- Train the models like kNN, Random Forest, SVM etc.
- Model evaluation using metrics like accuracy, precision and recall.
- Performance comparison of the respective models.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Methodology

- Reviewed literature on intrusion detection.
- Loaded and explored the NSL-KDD dataset.
- Data preprocessing, feature exploration and feature extraction using PCA for model development.
- Train the models like kNN, Random Forest, SVM etc.
- Model evaluation using metrics like accuracy, precision and recall.
- Performance comparison of the respective models.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Methodology

- Reviewed literature on intrusion detection.
- Loaded and explored the NSL-KDD dataset.
- Data preprocessing, feature exploration and feature extraction using PCA for model development.
- Train the models like kNN, Random Forest, SVM etc.
- Model evaluation using metrics like accuracy, precision and recall.
- Performance comparison of the respective models.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Methodology

- Reviewed literature on intrusion detection.
- Loaded and explored the NSL-KDD dataset.
- Data preprocessing, feature exploration and feature extraction using PCA for model development.
- Train the models like kNN, Random Forest, SVM etc.
- Model evaluation using metrics like accuracy, precision and recall.
- Performance comparison of the respective models.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Methodology

- Reviewed literature on intrusion detection.
- Loaded and explored the NSL-KDD dataset.
- Data preprocessing, feature exploration and feature extraction using PCA for model development.
- Train the models like kNN, Random Forest, SVM etc.
- Model evaluation using metrics like accuracy, precision and recall.
- Performance comparison of the respective models.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Methodology

- Reviewed literature on intrusion detection.
- Loaded and explored the NSL-KDD dataset.
- Data preprocessing, feature exploration and feature extraction using PCA for model development.
- Train the models like kNN, Random Forest, SVM etc.
- Model evaluation using metrics like accuracy, precision and recall.
- Performance comparison of the respective models.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Result

- Random Forest achieves highest accuracy of 99.87% in detecting intrusions. It is equivalent to the decision tree with max_depth parameter equal to 3.
- k-Nearest Neighbors with the value of n_neighbors parameter equal to 20 achieves 98.94%.
- Support Vector Machine achieves the accuracy of 97.04%.
- Gaussian Naive Bayes achieves the accuracy of 91.61%.
- Logistic Regression achieves the accuracy of 87.17%.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Result

- Random Forest achieves highest accuracy of 99.87% in detecting intrusions. It is equivalent to the decision tree with max_depth parameter equal to 3.
- k-Nearest Neighbors with the value of n_neighbors parameter equal to 20 achieves 98.94%.
- Support Vector Machine achieves the accuracy of 97.04%.
- Gaussian Naive Bayes achieves the accuracy of 91.61%.
- Logistic Regression achieves the accuracy of 87.17%.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Result

- Random Forest achieves highest accuracy of 99.87% in detecting intrusions. It is equivalent to the decision tree with max_depth parameter equal to 3.
- k-Nearest Neighbors with the value of n_neighbors parameter equal to 20 achieves 98.94%.
- Support Vector Machine achieves the accuracy of 97.04%.
- Gaussian Naive Bayes achieves the accuracy of 91.61%.
- Logistic Regression achieves the accuracy of 87.17%.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Result

- Random Forest achieves highest accuracy of 99.87% in detecting intrusions. It is equivalent to the decision tree with max_depth parameter equal to 3.
- k-Nearest Neighbors with the value of n_neighbors parameter equal to 20 achieves 98.94%.
- Support Vector Machine achieves the accuracy of 97.04%.
- Gaussian Naive Bayes achieves the accuracy of 91.61%.
- Logistic Regression achieves the accuracy of 87.17%.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

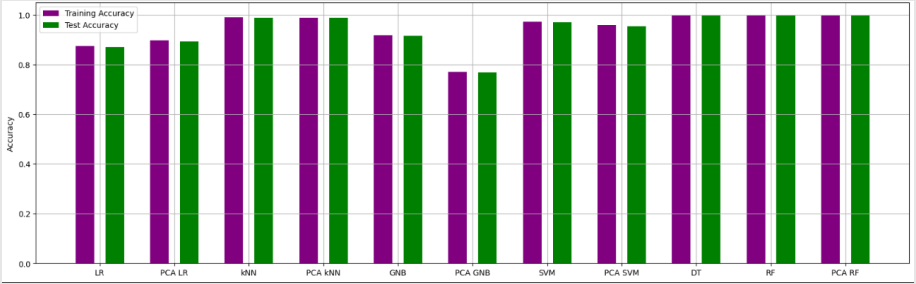
Result

- Random Forest achieves highest accuracy of 99.87% in detecting intrusions. It is equivalent to the decision tree with max_depth parameter equal to 3.
- k-Nearest Neighbors with the value of n_neighbors parameter equal to 20 achieves 98.94%.
- Support Vector Machine achieves the accuracy of 97.04%.
- Gaussian Naive Bayes achieves the accuracy of 91.61%.
- Logistic Regression achieves the accuracy of 87.17%.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Result (Continued)

The comparison between the respective results based on the accuracy is as follows:



विद्या ददाति विनयं विनयाद् याति पात्रताम्

Conclusion

- Machine learning algorithms like SVM, kNN, Random Forest are very important for intrusion detection for their capabilities of classification and prediction.
- The integration of AI and ML models enhance the ability to detect complex, unknown threats by learning patterns from large datasets, making intrusion detection systems more adaptive and intelligent over time.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Conclusion

- Machine learning algorithms like SVM, kNN, Random Forest are very important for intrusion detection for their capabilities of classification and prediction.
- The integration of AI and ML models enhance the ability to detect complex, unknown threats by learning patterns from large datasets, making intrusion detection systems more adaptive and intelligent over time.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Future Work

- In future, we will try to deploy the model for real-time intrusion detection using Kafka.
- We will also concentrate to use the Large Language Models for detection which can lead to automated response system.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Future Work

- In future, we will try to deploy the model for real-time intrusion detection using Kafka.
- We will also concentrate to use the Large Language Models for detection which can lead to automated response system.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

References

- Panwar, Shailesh Singh, and Y. P. Raiwani. "Data reduction techniques to analyze NSL-KDD Dataset." Int. J. Comput. Eng. Technol 5.10 (2014): 21-31.
- W. Li, P. Yi, Y. Wu, L. Pan, J. Li et al., "A new intrusion detection system based on knn classification algorithm in wireless sensor network," Journal of Electrical and Computer Engineering, vol. 2014, 2014.
- M. Karthikeyan, D. Manimegalai, and K. RajaGopal, "Firefly algorithm based wsn-iot security enhancement with machine learning for intrusion detection," Scientific Reports, vol. 14, no. 1, p. 231, 2024.

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Thank You!

विद्या ददाति विनयं विनयाद् याति पात्रताम्