Date: 27/04/2024

Project Fortnightly Report

Author: Aryan Deshmukh

Status of activities planned in [Current 15 Days/Last 15 Days]

Sr. No.	Activity	Duration	Status
1	• Research on Network Packet Transmission Data. • Study previous research papers on supervised machine learning techniques applied to network security. • Explore publicly available datasets related to network intrusion detection.	15 days	Complete
2	• Explore publicly available datasets related to network intrusion detection. • Clean the dataset by handling missing values and outliers. • Explore descriptive statistics and visualizations to gain insights into the data distribution. • Identify potential correlations between features and the target variable.	15 days	Complete
3	• Feature Selection and Engineering. • Conduct feature importance analysis to identify relevant features for classification. • Engineer new features based on domain knowledge and insights gained from data exploration.	15 days	Complete
4	• Model Selection and Training (KNN). • Evaluate the KNN algorithm's suitability for classification tasks. • Split the dataset into training and validation sets. • Train the KNN model and evaluate its performance metrics.	15 days	Complete

5	• Model Selection and Training (Decision Tree). • Evaluate the Decision Tree algorithm's suitability for classification tasks. • Split the dataset into training and validation sets. • Train the Decision Tree model and evaluate its performance metrics.	15 days	Complete
6	• Model Evaluation and Fine-Tuning. • Perform cross-validation to assess the generalization performance of models. • Fine-tune hyperparameters using techniques such as grid search or random search. • Compare the performance of tuned models and select the best-performing one for further analysis.	15 days	Complete
7	• Testing and Validation. • Evaluate the final models on a separate test dataset to assess their performance on unseen data. • Validate the models' effectiveness in distinguishing between normal and malicious network activity. • Generate performance metrics such as accuracy, precision, recall, and F1-score for model evaluation.	15 days	Complete

*Status can be [Complete] or [Extended to next 15 days] with reason