**Intrusion detection system for SDN**

* Dataset: NSL-KDD (train & test data set)
* Added column to dataset.
* Data Cleaning:
* There are not any null values in the dataset of train and test dataset.
* Data Pre-processing:
* Attack column in the dataset is merged to form attack\_categories. It is the target variable in the dataset.
* Classifying each attack to attack type. The different attack types are:
* Denial of Service attacks
* Probe attacks
* Privilege escalation attacks
* Remote access attacks
* Data Visualization:
* Bar plot for attack category.
* Analyze the target variable with features.
* Encoding:
* Changed the Object types in the train and test dataset and encoded them.
* Feature Selection:
* Heatmap is plotted.
* Finding features that are highly correlated (correlation coefficient > 0.5) with 'attack\_category'.
* Finding features that are low correlated (correlation coefficient < 0.1) with 'attack\_category'.
* Dropped the highly correlated and low correlated features with attack\_category.
* Feature Scaling.
* Standard Scaler is used for feature scaling, for normalizing the features.
* Model Training:
* SVM
* Overall Accuracy is 98%.
* However, accuracy alone may not be sufficient for evaluating the model’s performance.
* Confusion Matrix is plotted for more detailed information about the model’s performance.
* Overall, the SVM model demonstrates strong performance for most classes, with high precision, recall, and F1-scores. However, it struggles with class 0, where the recall is relatively low, indicating that the model misses a significant portion of the actual instances.
* ANN
* Overall, the ANN model demonstrates strong performance, with high accuracy and precision-recall values for most classes. However, like the SVM model, it struggles with class 0, where the recall is relatively low, indicating that the model misses a significant portion of the actual instances.