SIT719 Security and Privacy Issues in Analytics

Pass Task 4.1: Attack Classification using Naïve Bayes Algorithm

Overview

An Intrusion Detection System (IDS) is a system that monitors network traffic for suspicious activity and issues alerts when such activity is discovered. Supervised learning techniques have been proven very effective for intrusion detection.

This week, we have explored the supervised machine learnings and how to apply them for security solutions, e.g., intrusion detection in NSL-KDD dataset. We have demonstrated how classification algorithms can separate the normal instances from the attack classes.

In this pass task, we will employ “Naïve Bayes” algorithm. If you are interested to know how you can implement simple machine learning algorithms using WEKA, follow the link below:

https://www.youtube.com/watch?v=TF1yh5PKaqI&feature=emb_logo

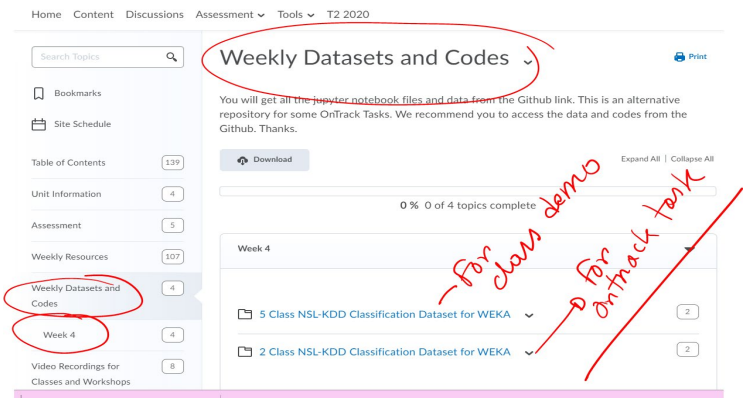
This is a *Pass task*, so you MUST complete the task and submit the evidence of your work to Ontrack.

Task Description

Instructions:

This task is a binary classification (2-class problem). Follow the below steps to complete the task. Once you have the results and reports, compile in a PDF and submit to the onTrack system.

1. Download the data folder from the CloudDeakin contents, look for Weekly Dataset and codes. Then look for 2-class NSL-KDD datasets (both train and test) and save in a folder.



Load the Train dataset into WEKA. Once uploaded, you may check the data distribution by selecting the class attribute and it will appear as Figure 2.

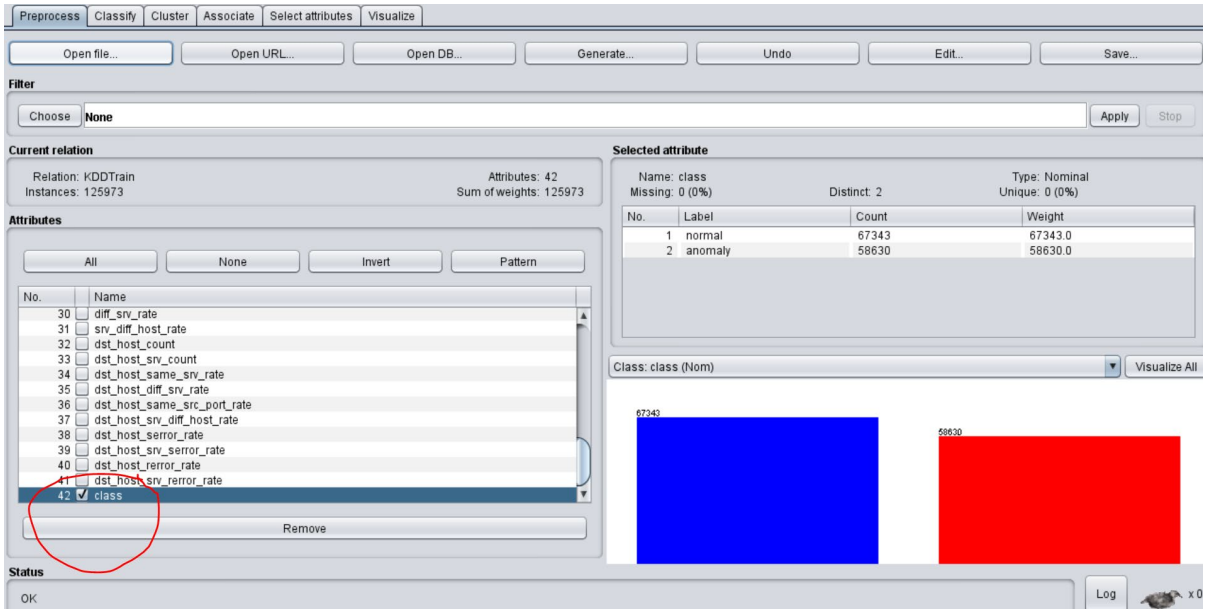


Figure 2

2. Now apply “Naïve Bayes” classification algorithm from the “Classify” tab.

3. Check the results with a 10-fold cross validation.

4. Now, upload the test dataset and check the classification results.

5. Compare the results between 10-fold cross validation and the one obtained using the test dataset. Use confusion matrix to explain the results.

Take screen shots for each of the steps (or results) and drop them into a word document. Finally, save the word document into a SINGLE PDF and upload to the OnTrack system.