TASK 2 – Intrusion Detector Learning (35%)

Software to detect network intrusions protects a computer network from unauthorized users, including perhaps insiders. The intrusion detector learning task is to build a predictive model (i.e. a classifier) capable of distinguishing between „bad“connections, called intrusions or attacks, and „good“ normal connections.

A connection is a sequence of TCP packets starting and ending at some well defined times, between which data flows to and from a source IP address to a target IP address under some well defined protocol. Each connection is labeled as either normal, or as an attack, with exactly one specific attack type. Each connection record consists of about 100 bytes.

Attacks fall into four main categories:

- DOS: denial-of-service, e.g. syn flood;
- R2L: unauthorized access from a remote machine, e.g. guessing password;
- U2R: unauthorized access to local superuser (root) privileges, e.g., various „buffer overflow“ attacks;
- probing: surveillance and other probing, e.g., port scanning.

In this second task, you will carry out analysis on a dataset gathered from such a network intrusion detection system, and you can find more general information about intrusion detector learning systems here.

Download the dataset here.

You will find two different sets, a training set, and a test set. To obtain comparable results among all participants, please make sure that you use the training set to develop your algorithm (see instructions below) and use the test set to evaluate your algorithm.

Your task here is to:

1. **Analyze the dataset to obtain helpful information about the data**.

2. **Based on your analysis, you should build a ML model that classfies *connections into different categories.***
   To pass this task, your classifier needs to differentiate between good/bad connections and needs to perform better than the majority classifier. Classifiers that can provide better accuracy and/or more fine-grained classifications (i.e., can determine the type of attack) will be rated higher. Your grade in this task is determined by the clarity of your analysis, how you link your analysis to your model, and, finally, the performance of your model. In this task, the grading criteria with regards to the classifier are the Detection Rate (for bad connections or different types of attacks), Precision, and Recall of your solution. Note that false positives and false negatives may incur different real-world costs on the classifier, and thus some metrics may be more important than others.

3. **You will present your findings in class. Please prepare presentation slides that illustrate your analysis findings, describe the way you have built your model, run a demo of your classifier and show the results of the prediction.**

4. **Also, at the end of the semester, a written report that visualizes your most interesting findings from the analysis (those which have impacted your model design) and illustrates the performance of your model ist o be submitted. At this point, also submit your code.**

   Note: This dataset was published in: Stolfo, J., et al. "Cost-based modeling and evaluation for data mining with application to fraud and intrusion detection." Results from the JAM Project by Salvatore (2000).