Network Intrusions clustering

Anomalies, also referred to as outliers, are defined as a set of patterns which significantly deviate from the normal or expected pattern. In our digital age, anomalies are an integral part of many applications including cyber security , manufacturing, fraud detection, health-care systems and numerous other fields. For instance, in cyber security, intrusion detection system can identify an anomalous pattern like unauthorized access to sensitive information or security violation. One of the major challenges include the scarcity of anomalous labelled data. Most of the systems do not have or few anomalous patterns and make the learning task difficult. To solve this situation, unsupervised anomaly detection model has the ability to learn the model without anomalous data patterns.

There are already high performing proposed solutions for anomaly binary classification in the domain of cyber security. To better design a defense system it is important to identify the types or clusters of the anomalies. In this task you will be using an unsupervised approach to cluster network intrusions. The dataset for the task can be found **here**. All the steps for clustering of intrusions should be clearly outlined in the notebook or python script. Use machine learning method of your choice. Reason of choosing the method should be stated in the notebook or python script