**ENGINEERED FEATURES – ANOMALY DETECTION**

1. **Connection-based features**: Rolling window for the 1000 previous packets when **SRCADDRESS** appears in the traffic.

* **How many netflows exist in a given time frame?**
* **How many times the SRCADDRESS has appeared in the last 100 netflows**
* For any of the flow records that SRCADDRESS has appeared within the last 1000 flow records , count the distinct destination ports
* For any of the flow records that SRCADDRESS has appeared within the last 1000 flow records , count the distinct source ips
* For any of the flow records that SRCADDRESS has appeared within the last 1000 flow records , count the distinct source ports
* For any of the flow records that SRCADDRESS has appeared within the last 1000 flow records , average the packets
* For any of the flow records that SRCADDRESS has appeared within the last 1000 flow records , average the bytes
* **Repeat Features for DSTADDRESS**

1. **Time-based features**: Rolling window for the 10 previous minutes when **SRCADDRESS** appears in the traffic.

* How many records exist in a given time frame?
* How many times has SRCADDRESS has appeared within the last 10 minutes
* For any of the flow records that SRCADDRESS has appeared within the last 10 minutes, count the distinct source IPs (SRCADDRESS)
* For any of the flow records that SRCADDRESS has appeared within the last 10 minutes, count the distinct source ports (SRCPORT)
* For any of the flow records that SRCADDRESS has appeared within the last 10 minutes, get the average packets
* For any of the flow records that SRCADDRESS has appeared within the last 10 minutes, get the average bytes
* **Repeat Features for DSTADDRESS**

**Each feature will result into a new column added.**