Project proposal

Network intrusion detection:

This is proposed as a requirement of T5 Bootcamb.

Introduction:

The Intrusion Detection System is based on monitoring the flow of data within the network and giving an alert either through behavior, connection, or host

We have systems that detect the movement of malicious data when it enters the network and these systems exist in **two types** (IDS and IPS)

IDS: It monitors If the network detects malicious data traffic, it sends a warning to the network engineer or protection official in order to counter the malicious data traffi

IPS: It detects malicious data traffic after it is detected and blocks it with the protection officer announcing that.

Purpose of the study:

The study aims to detect network intrusion to add security solutions to other security solutions to provide protection for the network.

We have two types of application for intrusion detection systems. I hope the application is at the network level, where malicious activities that occur within the network are detected, or the application is at the host level, where malicious activities are detected at the level of one device

Dataset:

This dataset can be found at Kaggle.

The dataset to be audited was provided which consists of a wide variety of intrusions simulated in a military network environment.

data shape:(25193,10)

- **Duration**
- **protocol type (**tcp , udp**)**
- **service (**http, private**)**
- **flag (**SF,S0)
- **src_bytes**
- **dst_bytes**
- **land**
- **wrong_fragment**
- **urgent**
- **hot**

duration	protocol type	service	flag	src_bytes	dst_bytes	land	wrong_fragment	urgent	hot
0	tcp	ftp_data	SF	491	0	0	0	0	0

QS on the DS :

1- if UDP or TCP can affect the hacker
2- the predict to attack is benign or harmful
3- what type service
4- type connection is normal or specific attack type

Tools:

There are different tools we would work with as Machin Learning:

- Jupiter notebook
- Pandas
- NumPy
- Matplotlib

TO DO:

- Explore the data and come up with EDA phases then use a model to fit the data.
- NOTE: the used features may be increased or changed and the model as well.