* **Cyber Attacks analysis using machine learning techniques**

**Software to detect network intrusions protects a computer network from unauthorized users.  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.**

**This project aims at analyzing the cyber attacks data set by using Gaussian Naive Bayes Method and Logistic regession method.The correlated columns of the data set are identified by generating the heat map, and those columns are dropped to reduce the unnecessary load on the data set.**

**The project comprises of four modules. The first module deals with the identification of correlated columns by generating the heatmap. The second module deals with the Gaussian Naive Bayes analysis.The third module deals with the Logistic Regression analysis.The fourth module deals with the generation of a comparison plot between the Gaussian Naive Bayes and logistic regression.**

**A Naive Bayes classifier is a probabilistic machine learning model that’s used for classification task by using Baye's theorem of probability.Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem.A closed-form expression is a mathematical expression that can be evaluated in a finite number of operations. It may contain constants, variables, certain "well-known operations and functions. Maximum-likelihood training can be done by evaluating a closed-form expression, which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.**

**Tools Used: Sklearn’s Gauusian Naive Bayes & Logistic Tregression**

**OutPuts from the project:**

**(1) Heatmap plot**

**(2) Gaussian Naive Bayes Accuracy Score**

**(3) Logistic Regression Score**

**(4) Comparison Plot**