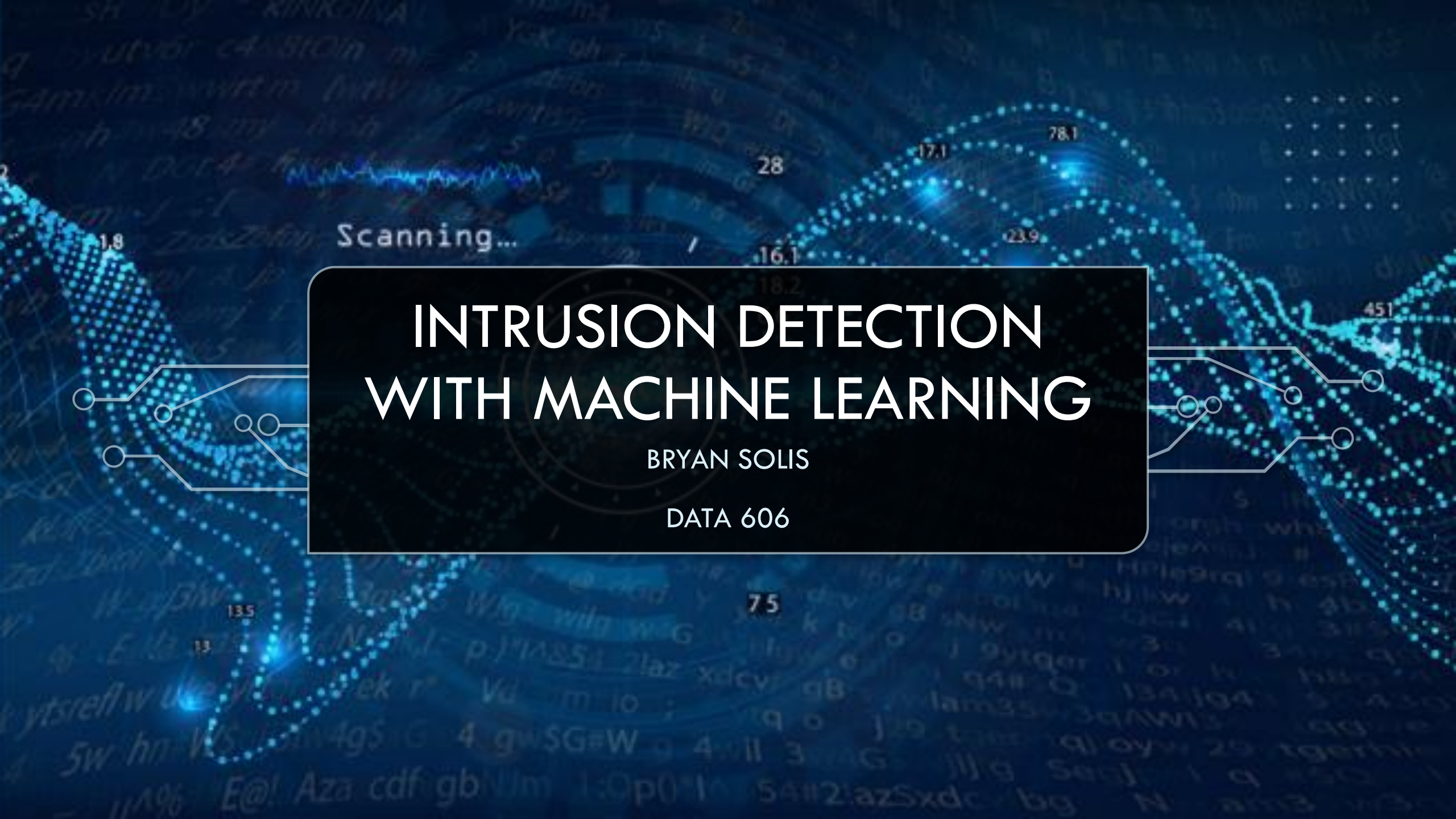


INTRUSION DETECTION WITH MACHINE LEARNING

BRYAN SOLIS

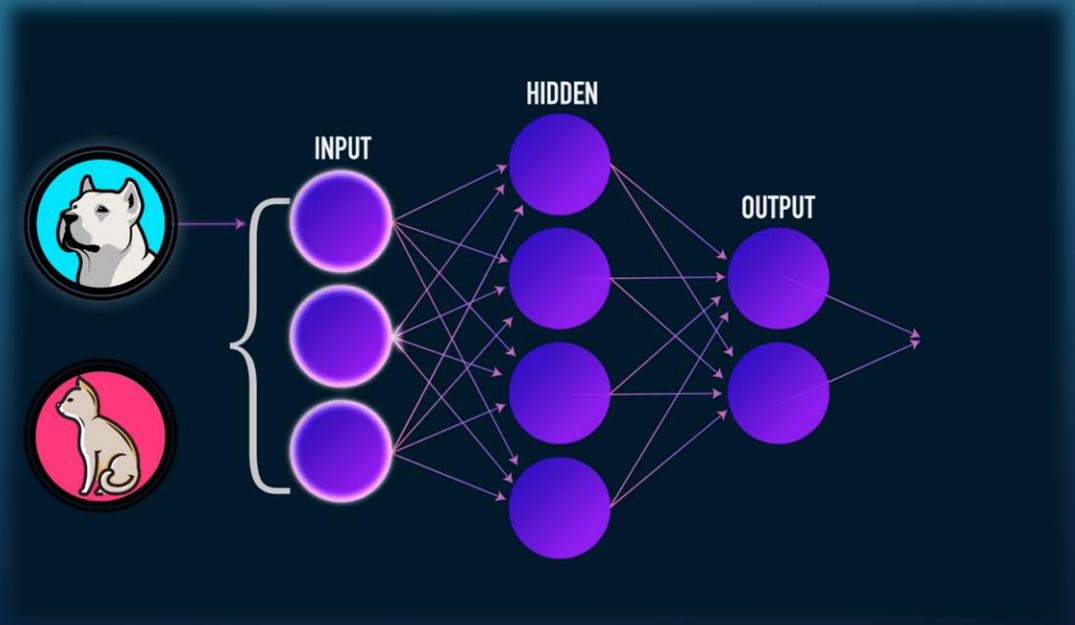
DATA 606

WHAT IS INTRUSION DETECTION?

- It aims to monitor and identify any unusual access or attack in the system or network.
- It's crucial for network security because it enables you to detect and respond to malicious traffic.
- The primary benefit of an intrusion detection system is to ensure IT personnel is notified when an attack or network intrusion might be taking place.

PROJECT DESCRIPTION & GOAL

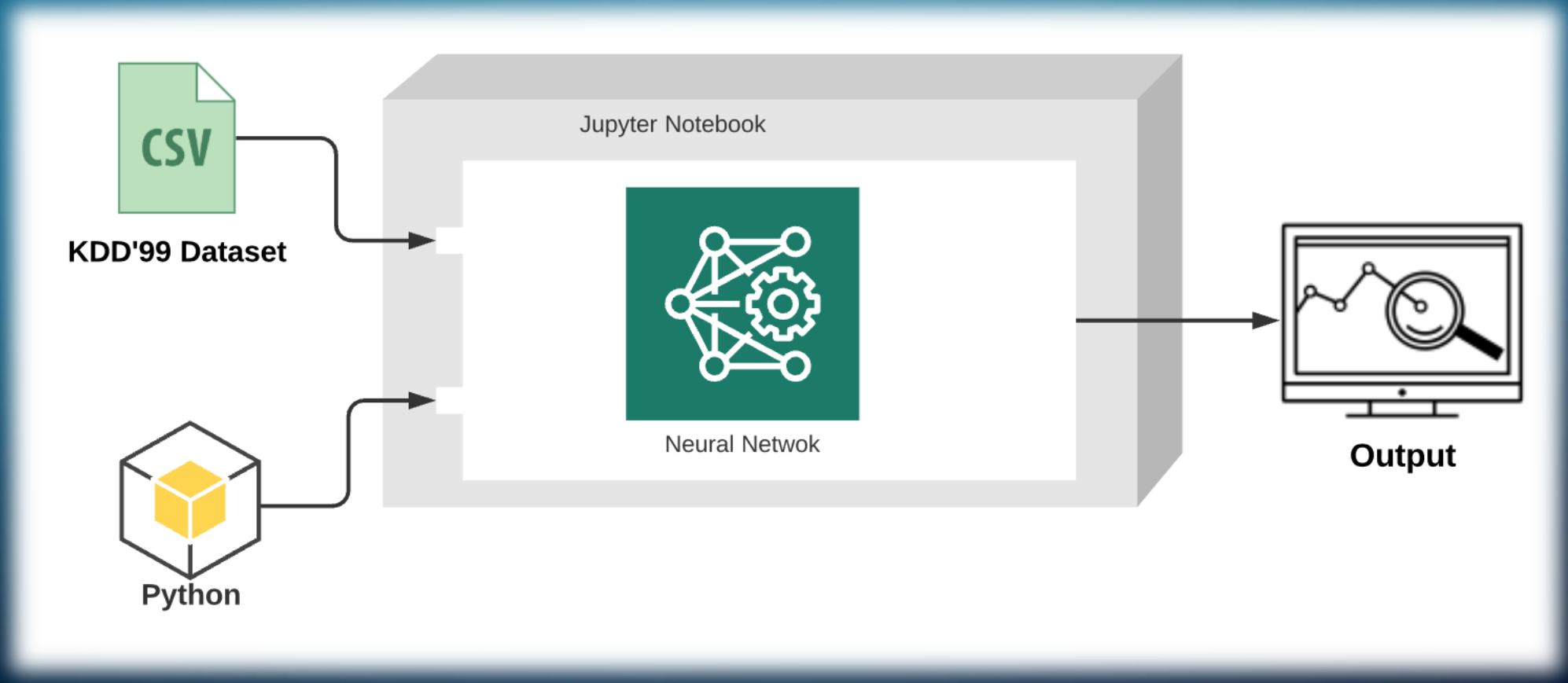
- Network traffic produce so much data that we need to find a way to properly predict types of attack against a network.
- By using Neural Network, this project will analyze traffic and will determine the type of attack that is being conducted in the network.
  - Build a model that can predict malicious traffic
  - If time allows, we will create a second model using Random Forest



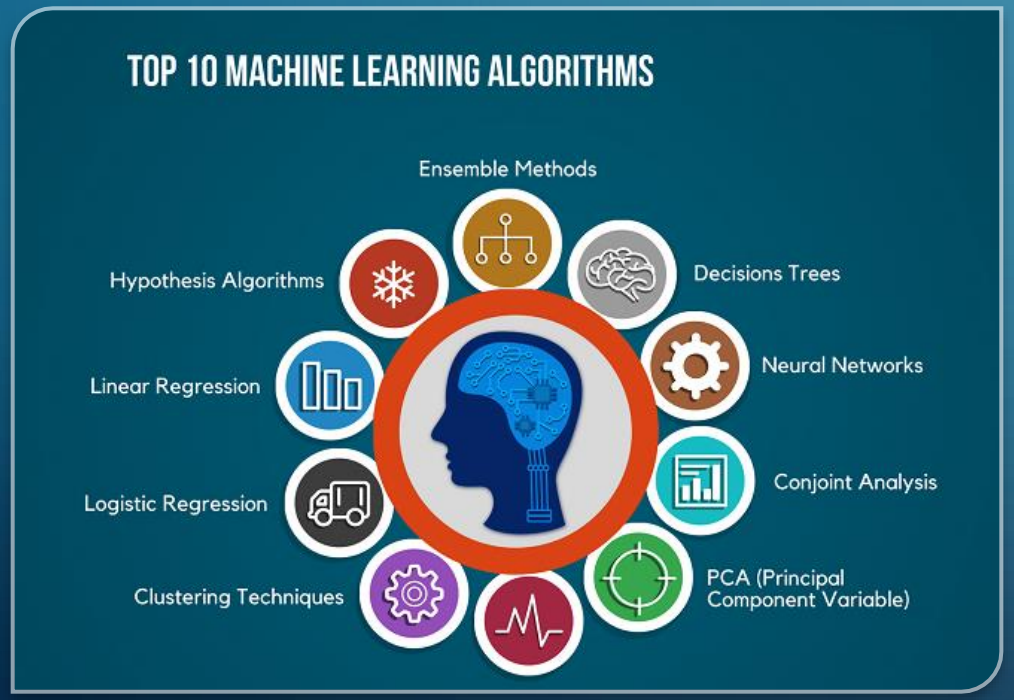
KDD'99 DATASET

- KDD'99 has been the most popular data set used for the evaluation of anomaly detection methods.
- This data set is prepared by Stolfo et al. and is built based on the data captured in DARPA'98 IDS evaluation program.
- It contain 41 unique features including label
- The data set is 75MB CSV File

ARCHITECTURE

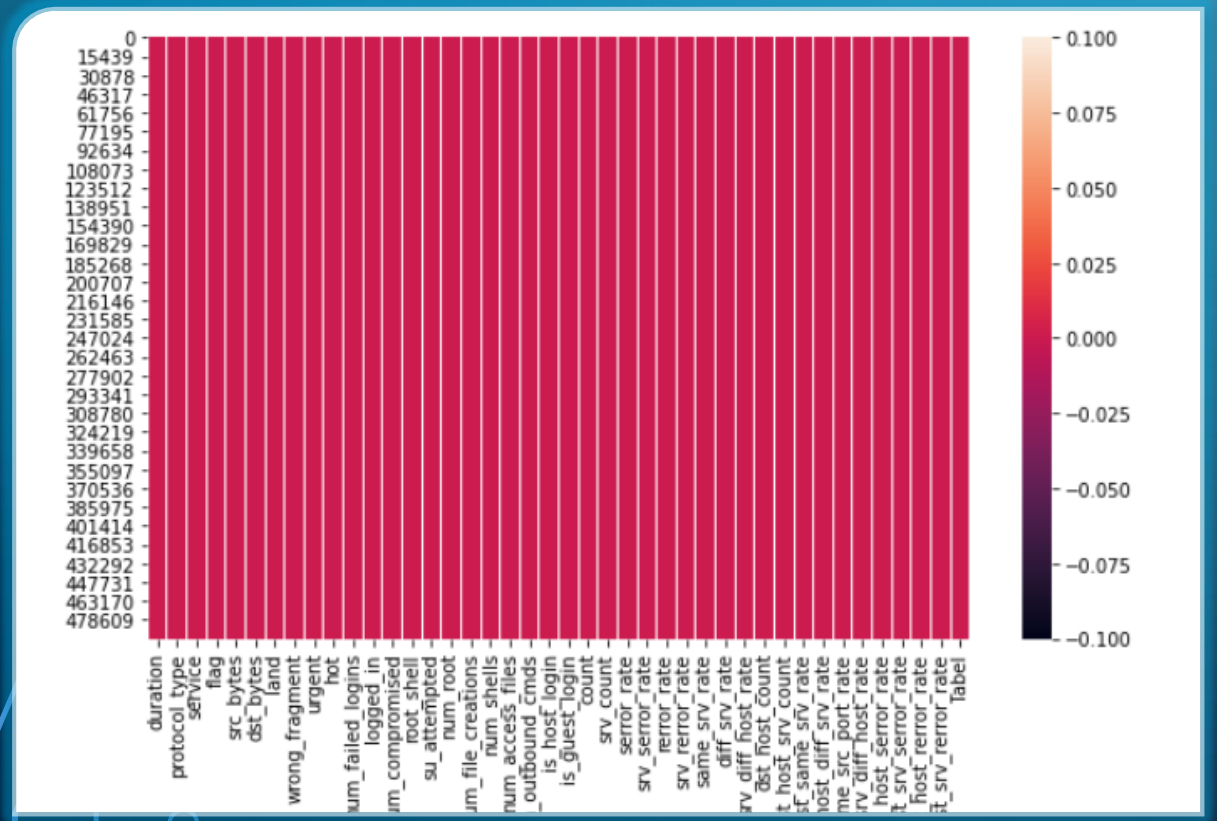


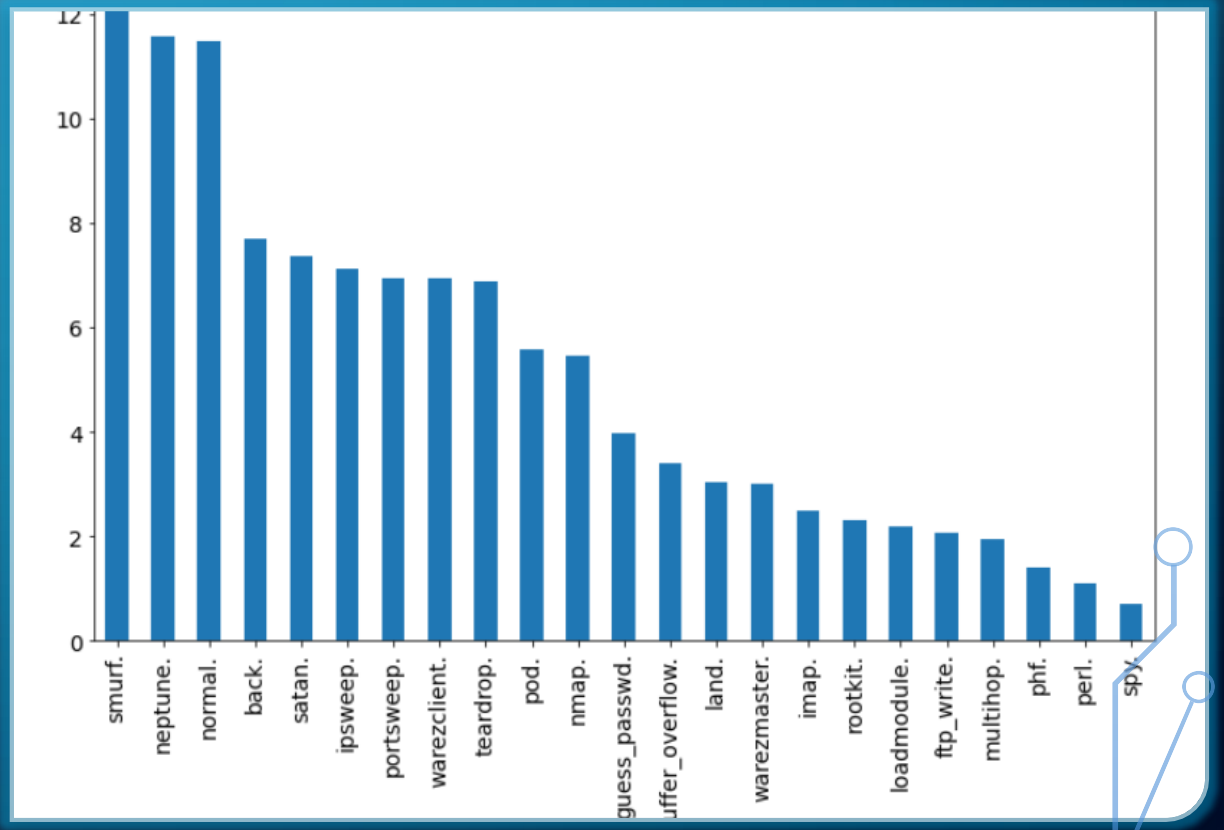
RELATED WORKS



- Similar work has been done using two different models
  - Spark Chi SVM. The author uses ChiSqSelector for features selection, and building a model using Support Vector Machine (SVM).
  - K-Means. The author uses Mini Batch K-means combined with principal component analysis (PCA).

DATA EXPLORATION





QUESTIONS ?


WHY?
WHAT?
WHERE?
WHEN?
HOW?
WHO?
WHICH?
WHOSE?
WHAT? HOW? WHY?
WHO? WHERE? WHAT? HOW?
WHY? HOW? WHERE?
WHO? WHOSE?
WHERE? WHAT? HOW?
WHERE? WHICH? WHOSE? WHEN? WHY?
HOW? WHO?
WHO? WHERE? WHAT? HOW?
WHAT? HOW? WHY?

REFERENCES

- Tsai, C., Hsu, Y., Lin, C., & Lin, W. (2009, May 20). Intrusion detection by machine learning: A review. Retrieved September 23, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0957417409004801
- Othman, S.M., Ba-Alwi, F.M., Alsohybe, N.T. *et al.* Intrusion detection model using machine learning algorithm on Big Data environment. *J Big Data* 5, 34 (2018). https://doi.org/10.1186/s40537-018-0145-4
- Ferhat K, Sevcan A. Big Data: controlling fraud by using machine learning libraries on Spark. Int J Appl Math Electron Comput. 2018;6(1):1–5.
- S. J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. Chan, "Costbased modeling for fraud and intrusion detection: Results from the jam project," discex, vol. 02, p. 1130, 2000.
- R. P. Lippmann, D. J. Fried, I. Graf, J. W. Haines, K. R. Kendall, D. McClung, D. Weber, S. E. Webster, D. Wyschogrod, R. K. Cunningham, and M. A. Zissman, "Evaluating intrusion detection systems: The 1998 darpa off-line intrusion detection evaluation," discex, vol. 02, p. 1012, 2000
- Kumar, V. (2019, November 21). Why Are Neural Networks Not the Answer to Everything? Retrieved September 24, 2020, from https://www.analyticsinsight.net/neural-networks-not-answer-everything/