Modelling & Predicting Cyber Hacking Breaches by ML

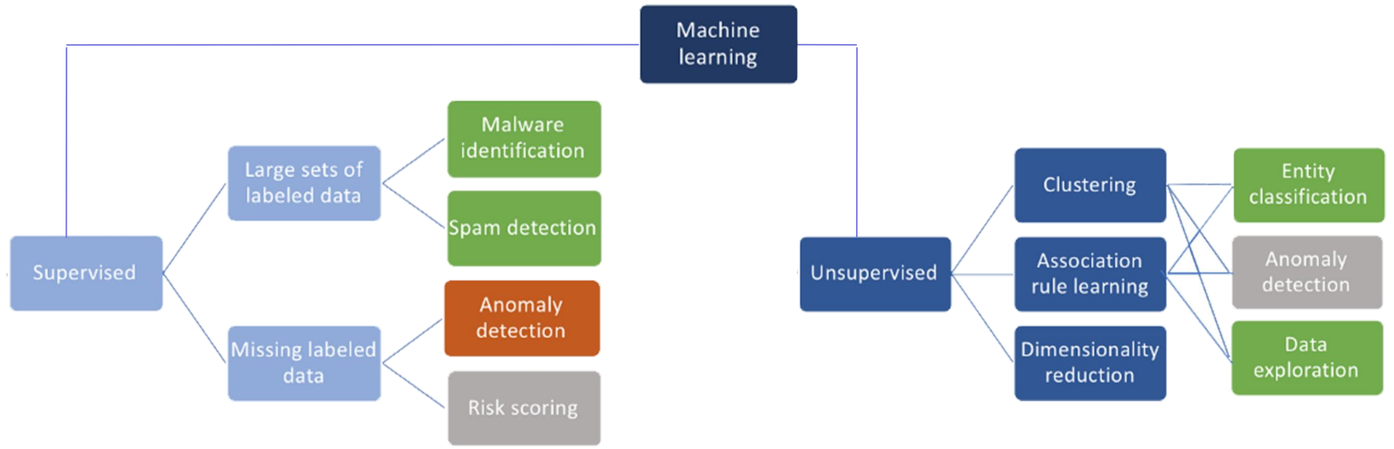
Bhushan Mahajan
MSc in Big Data

UNIVERSITY of STIRLING

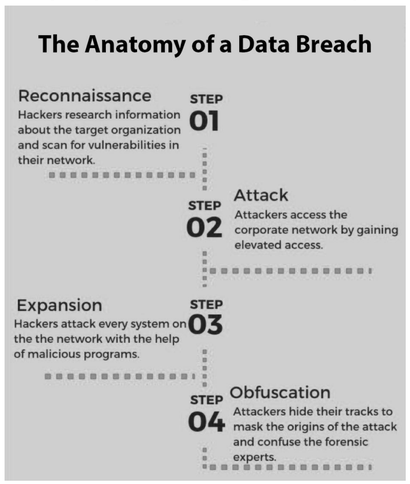
Introduction

Computers in the internet have given way to some of the biggest data breaches of all time- it's alarming!!! because these technologies are crucial to our way of life. We read the news on it check our email look at the weather and review our bank statements, they are so essential the last thing we want to think about is the possibility of them being an unstable security nightmare. Despite increases in security hackers always seem to find a way in or around the measures, from big-box retailers to bank no target is too big or small for them and the cost is enormous, resulting in losses reaching the millions ready to get a little paranoid about your passwords and bank accounts.

A cyber hacking breach occurs when a cybercriminal successfully infiltrates a data source and extracts sensitive information. This can be done physically by accessing a computer or network to steal local files or by bypassing network security remotely. The latter is often the method used to target companies.



Machine Learning in Data Breach



Scope And Objectives

In this study developing/building of models using machine learning techniques which can predicts future cyber hacking breach incidents so this can be the helpful techniques for companies to be already aware and ready for future incidents and to fulfill their weakness for the future happening incidents.

- Identify and narrate about data beach in between timing for cyber hacking braches.
- What was the breach size and approach used by criminals for hacking.
- Was the in between time/size for data breach dependant/independent of each other.
- Build a model which can predict the future incidents, what would be breach size and trend exhibit.
- Investigate how to predict the extremely large values and how to deal with missing data.

State of the Art

Latest and most current condition For modeling and predicting data breaches is Statistical analysis was published by authors. They proposed using a copula-based approach to predict the joint probability that an incident with a certain magnitude of breach size will occur during a future period of time. But statistical analysis is for inference about the relationships between variables. And this study will going to present a Machine learning to make the most accurate predictions possible.

Achievement

- Feature Extraction
- Hadns on Approach
- Algorithmic Methods

For further information contact: *Bhushan Mahajan* *bhm00006@students.stir.ac.uk*