

# Advancement of Machine Learning For Data Security

Anish Mahapatra | 21 February, 2021

# Agenda

---

- 01 Introduction
- 02 What is machine learning?
- 03 Components of Data Security
- 04 Case Study: ML for Data Security
- 05 Machine Learning for Data Security
- 06 Tools & Technologies

# Introduction




Senior Data Scientist 🧐

 **LinkedIn:** [www.linkedin.com/in/anishmahapatra](https://www.linkedin.com/in/anishmahapatra)

 **Medium:** [anishmahapatra.medium.com](https://anishmahapatra.medium.com)

 **Website:** [www.anishmahapatra.com](https://www.anishmahapatra.com)

 **GitHub:** [www.github.com/anishmahapatra](https://www.github.com/anishmahapatra)

Q. How to get into Data Science?

A. Simple, ask a real-world Data Scientist.



# What is Machine Learning?

---

“Machine learning preemptively stamps out cyber threats and bolsters security infrastructure through pattern detection, real-time cyber crime mapping and thorough penetration testing”

# What is Machine Learning?

01

## Data Access

Identify List of Data Owners,  
Request Access,  
Maintain Single POC,  
Understand current process

02

## Data Silos to Data Warehouse

Make Data Dictionary,  
Set up Data pipelines,  
Make ER Diagram,  
Main Analytical Data frame

03

## Data Analytics & Modelling

Exploratory Data Analysis,  
Business Analytics,  
Heuristic, Statistical & ML Models.  
Experimentation

06

## Business Decisions, Impact Analysis

Model Evaluation,  
Monitoring Model Performance and Drift,  
Data Governance Policies

05

## Data Visualization & Dashboarding

Dashboards to aid decision-making,  
Decide concurrency, cadence and security

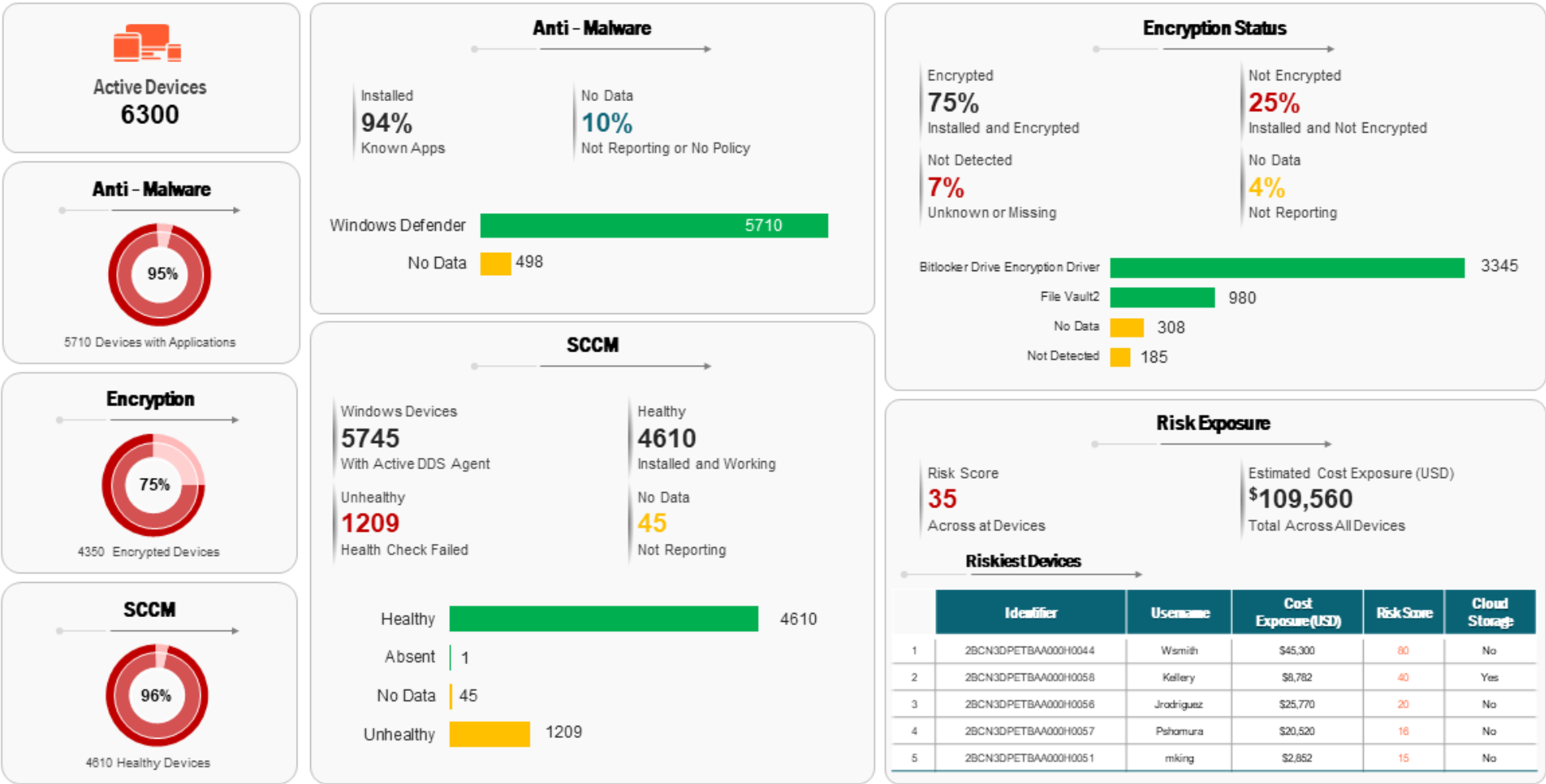
04

## Feature Engineering

Building new features,  
Combining external data,  
Building a feature store

# Security Dashboard

## Business Data Security Dashboard of Critical Application



# Components of Data Security

01



## Confidentiality

Privacy to the sensitive information while it is in transit over a network

02



## Integrity

Refers to preventing data from being tampered with, modified, or altered in an unauthorized way to achieve malicious goals

03

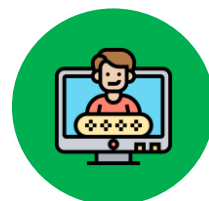


## Availability

Security service which ensures the constant availability of resources and services to only authorized parties in a timely manner.

### CIA Triad

04



## Authenticity

Refers to the state of being genuine, verifiable or trustable.

05

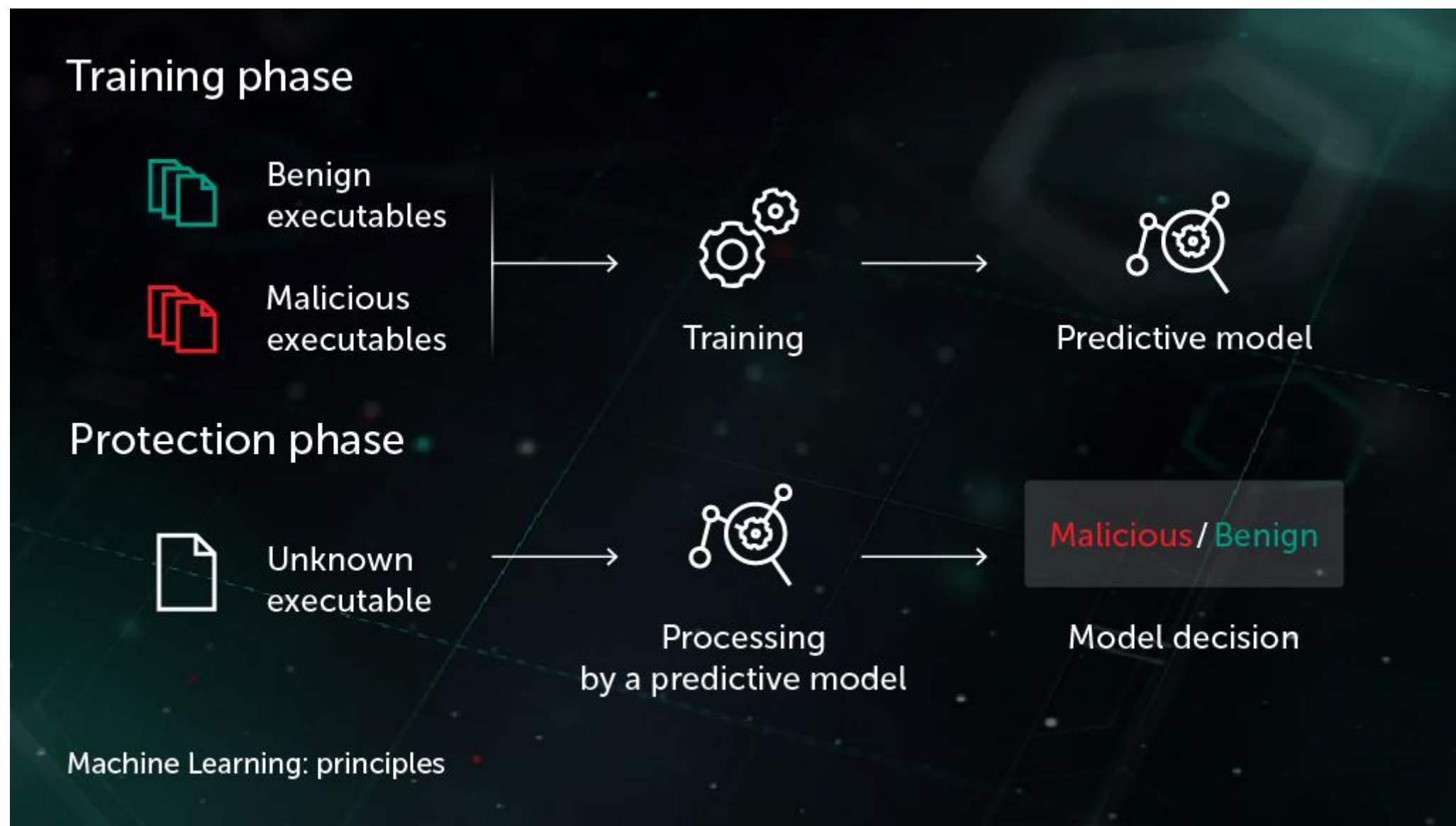


## Accountability

Refers to the ability to trace back the actions to the entity that is responsible for them

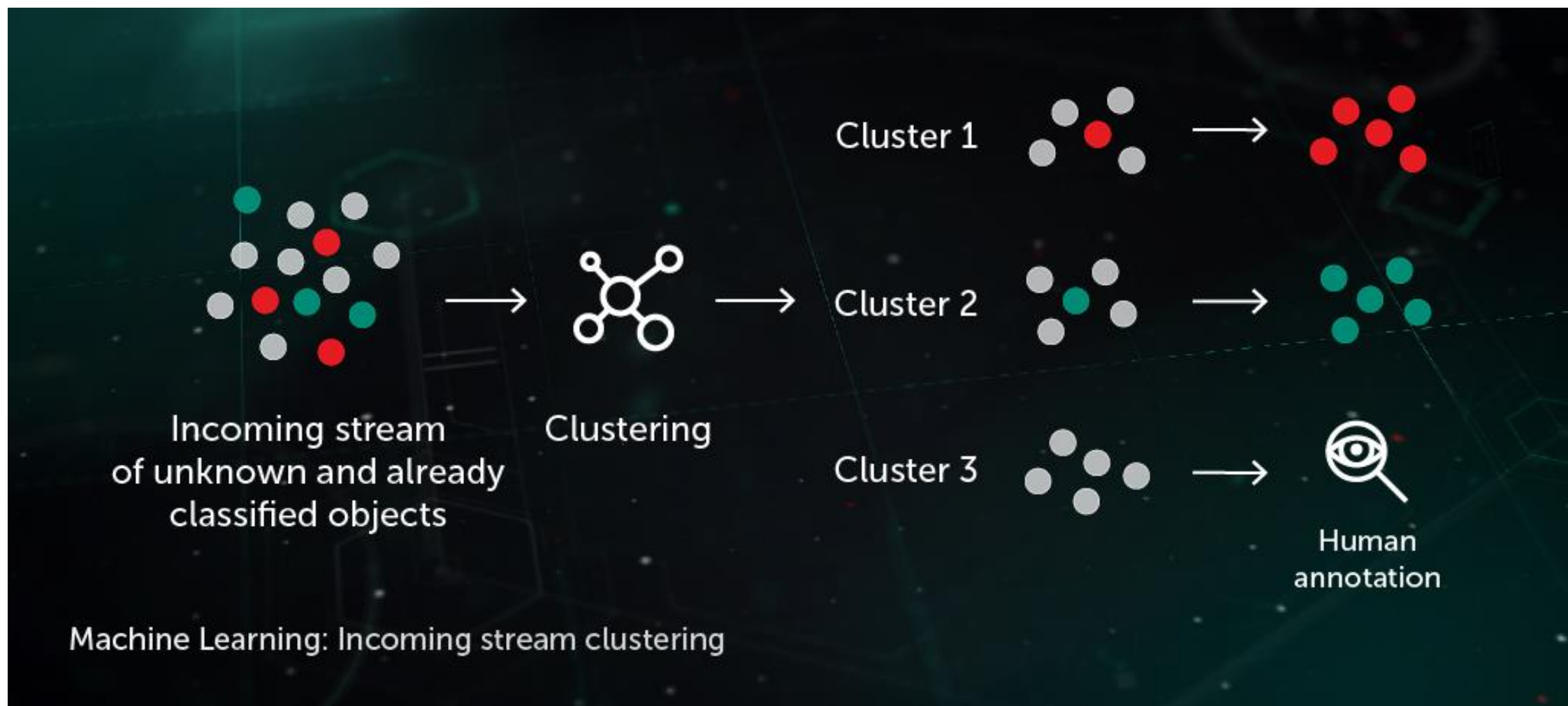


# Case Study: ML for Data Security - I





# Case Study: ML for Data Security - I



# Case Study: ML for Data Security - II

## Cyber Security For Attacks

Sub Class Name ▼

Select date Range ▼

Threats

88.8 K

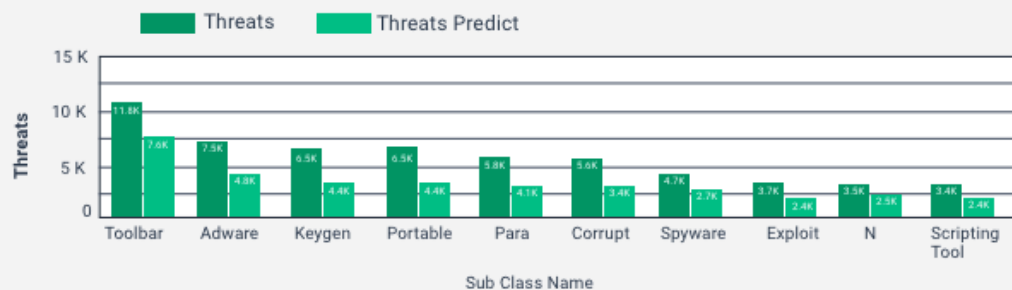
Phishing Attack

19.6 K

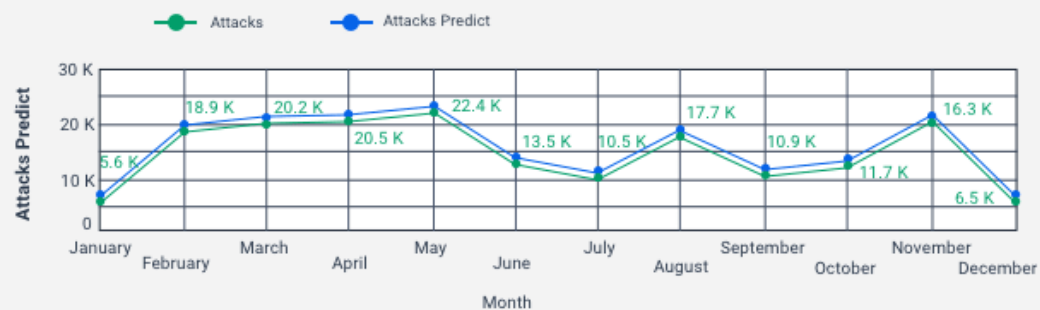
Attacks

174.8 K

### Top Threats By Class Name



### Attacks by Month



## Cyber Security

Select date Range ▼

Complaints

61.8 K

Non-Complaints

49.3 K

Open-Ticket

4.6 K

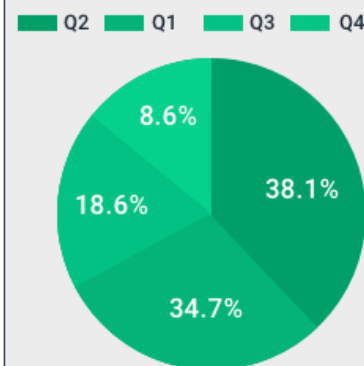
Close-Ticket

4.6 K

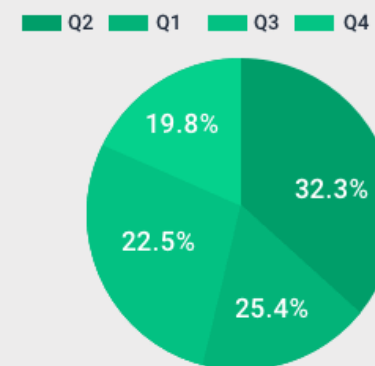
### Open Ticket by Month



### Complaints by Quarter



### Security Incident by Quarter



# Machine Learning for Data Security

01



## Network Threats

Monitor Network for Anomalies, ML allows us to identify detection of insider threats, unknown malware, policy violations

02



## Safe Browsing

ML can predict “bad neighborhoods” by analyzing internet activity. It identifies attack infrastructures

03



## Endpoint Malware protection

Algorithms can detect never-before-seen malware attacks that is trying to run on the endpoints

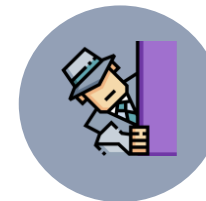
04



## Protect Data in the Cloud

Machine Learning can identify suspicious cloud app login activity, detecting location-based anomalies, and conducting IP reputation analysis

05



## Detect Malware in encrypted traffic

Analyze encrypted traffic data elements in common network telemetry. Malicious patterns can be identified without decrypting

# Tools & Technologies in Data Science



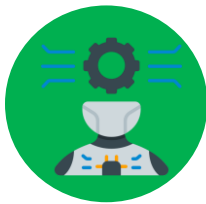
## Data Platforms



## Languages / Development Tools



## Data Exploration / Visualization



## ML / AI Tools



# Key Takeaways

---

- 01 Introduction
- 02 What is machine learning?
- 03 Components of Data Security
- 04 Case Study: ML for Data Security
- 05 Machine Learning for Data Security
- 06 Tools & Technologies



# Advancement of Machine Learning for Data Security

Anish Mahapatra

## Q&A Session



LinkedIn: [www.linkedin.com/in/anishmahapatra](https://www.linkedin.com/in/anishmahapatra)

Email: [anishmahapatra01@gmail.com](mailto:anishmahapatra01@gmail.com)

# Introduction




## Senior Data Scientist 🧐

 **LinkedIn:** [www.linkedin.com/in/anishmahapatra](https://www.linkedin.com/in/anishmahapatra)

 **Medium:** [anishmahapatra.medium.com](https://anishmahapatra.medium.com)

 **Website:** [www.anishmahapatra.com](https://www.anishmahapatra.com)

 **GitHub:** [www.github.com/anishmahapatra](https://www.github.com/anishmahapatra)

Q. How to get into Data Science?

A. Simple, ask a real-world Data Scientist.





# Advancement of Machine Learning For Data Security

Anish Mahapatra | 21 February, 2021

*Q. How to get into Data Science?*

A. Simple, ask a real-world Data Scientist.

# Thank you.

Email: [anishmahapatra01@gmail.com](mailto:anishmahapatra01@gmail.com)