Microsoft

Microsoft Security Incident Prediction

Microsoft

Microsoft's GUIDE Dataset

13 million data points across **33 entity** types, covering **1.6 million** alerts and **1 million** annotated incidents, from over **6,100 organizations**.

Microsoft

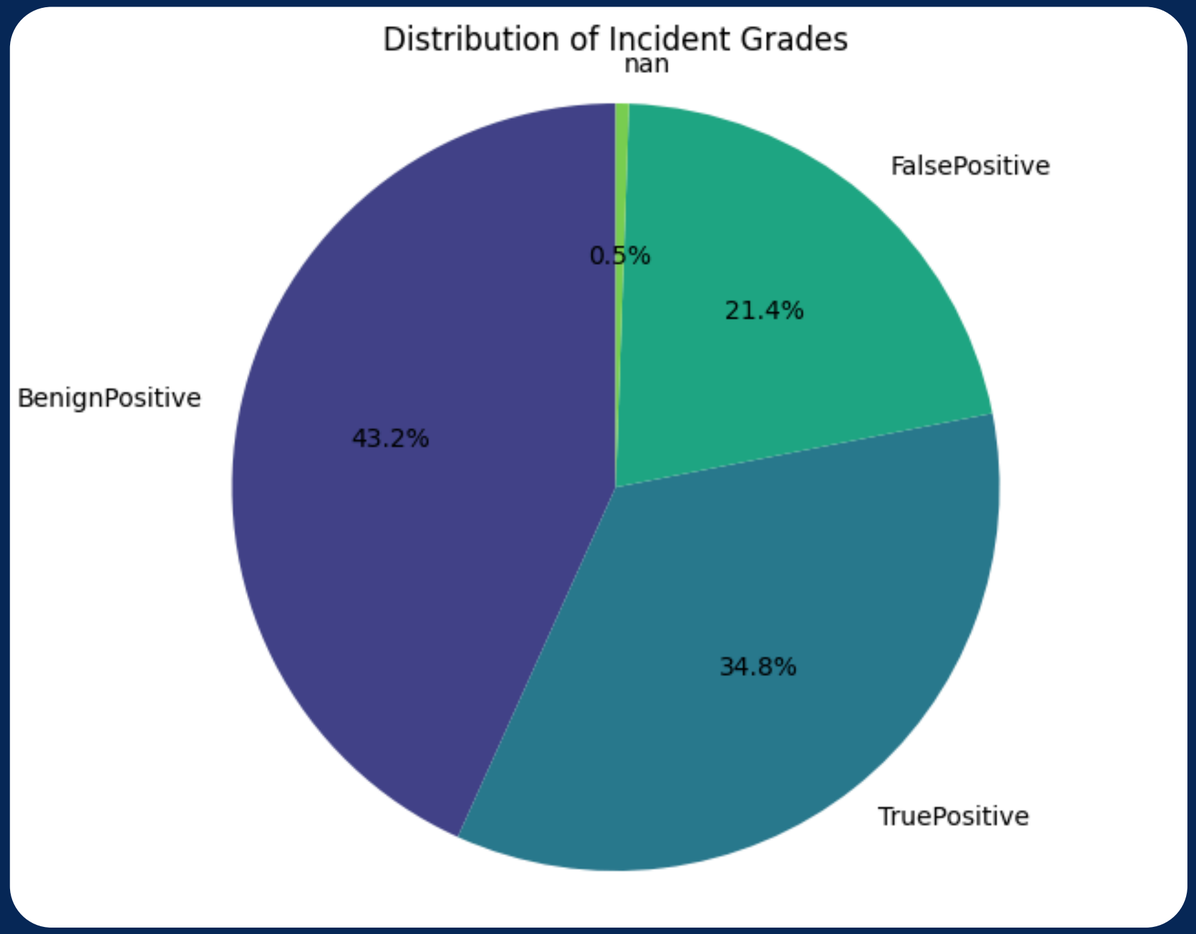
Agenda

- Data Exploration
- Data Analysis
- Automation in Incident Detection
- Top MITRE Attack
- Mitigation Techniques and Tools
- Future Directions and Conclusion

Microsoft

DATA EXPLORATION

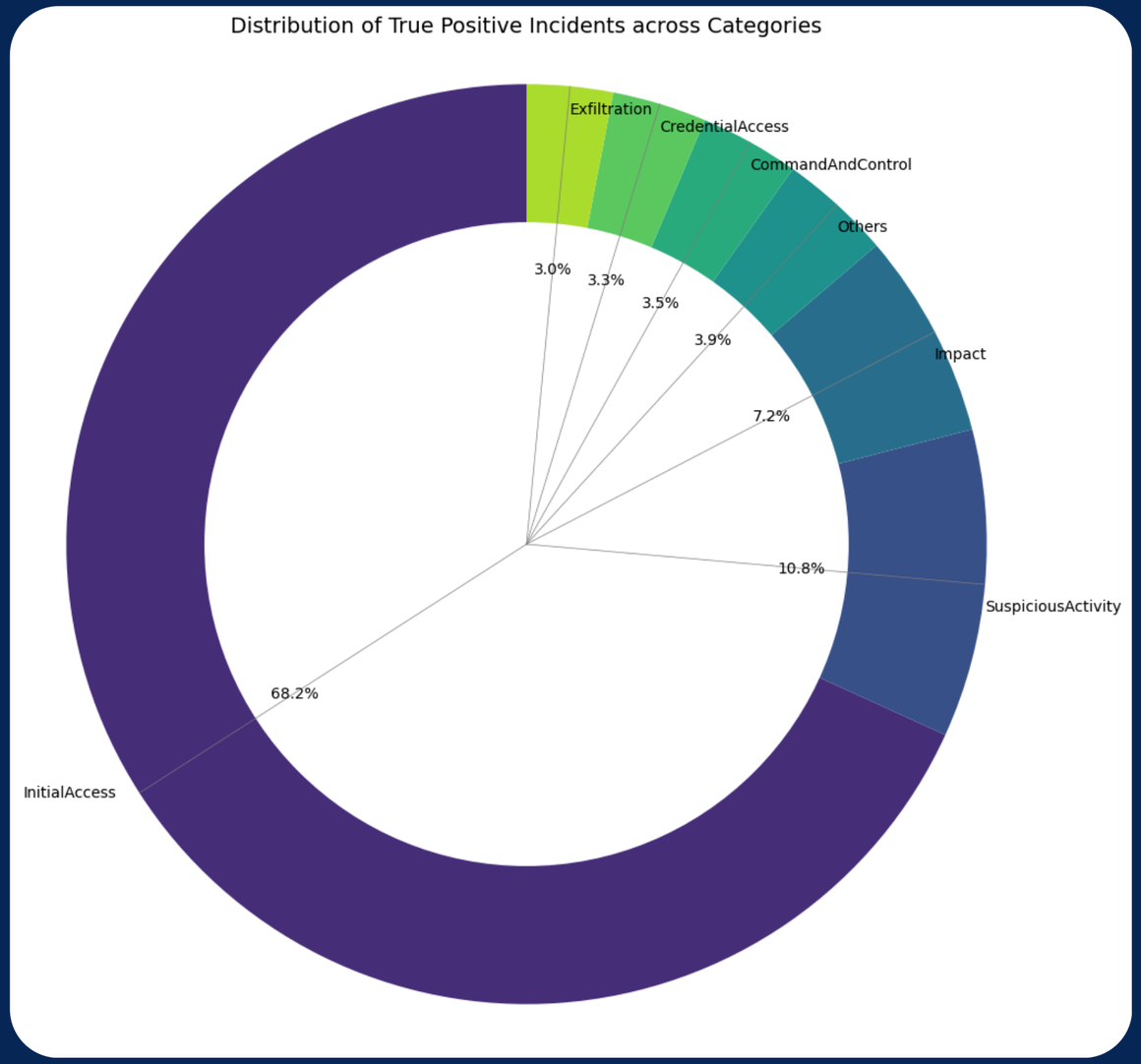
What Truly Matters?



35%

35% of the total detected incidents are classified as True Positives.
Indicating that they require immediate attention and action.

Different Categories of Incident



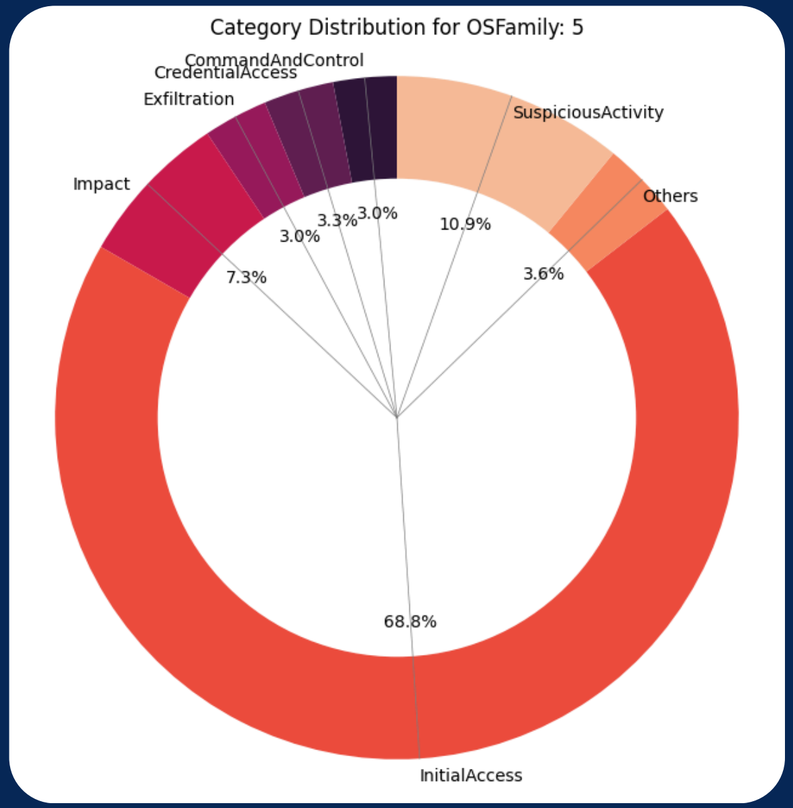
Initial Access 70%

Suspicious Activity 10.8%

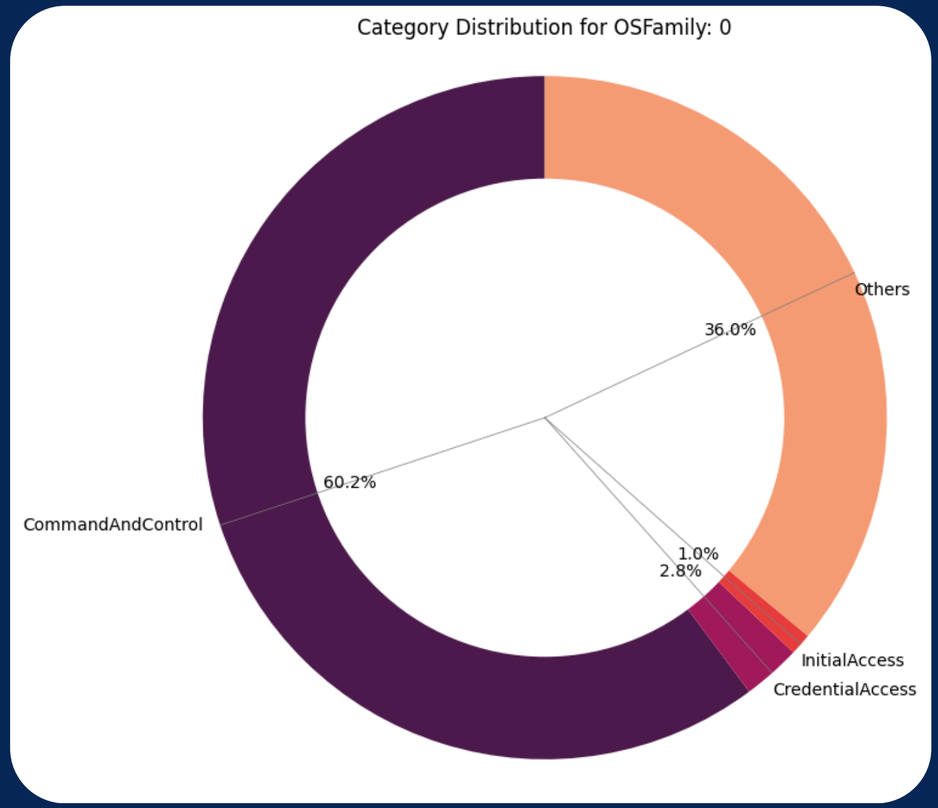
Impact 7.2%

Different Operation Systems

OSFamily	Proportion
5 - Windows	90.88%
0 - Linux	7.55%
1 - MacOS	1.51%
2 - Unix	0.05%
4 - Android	0.02%

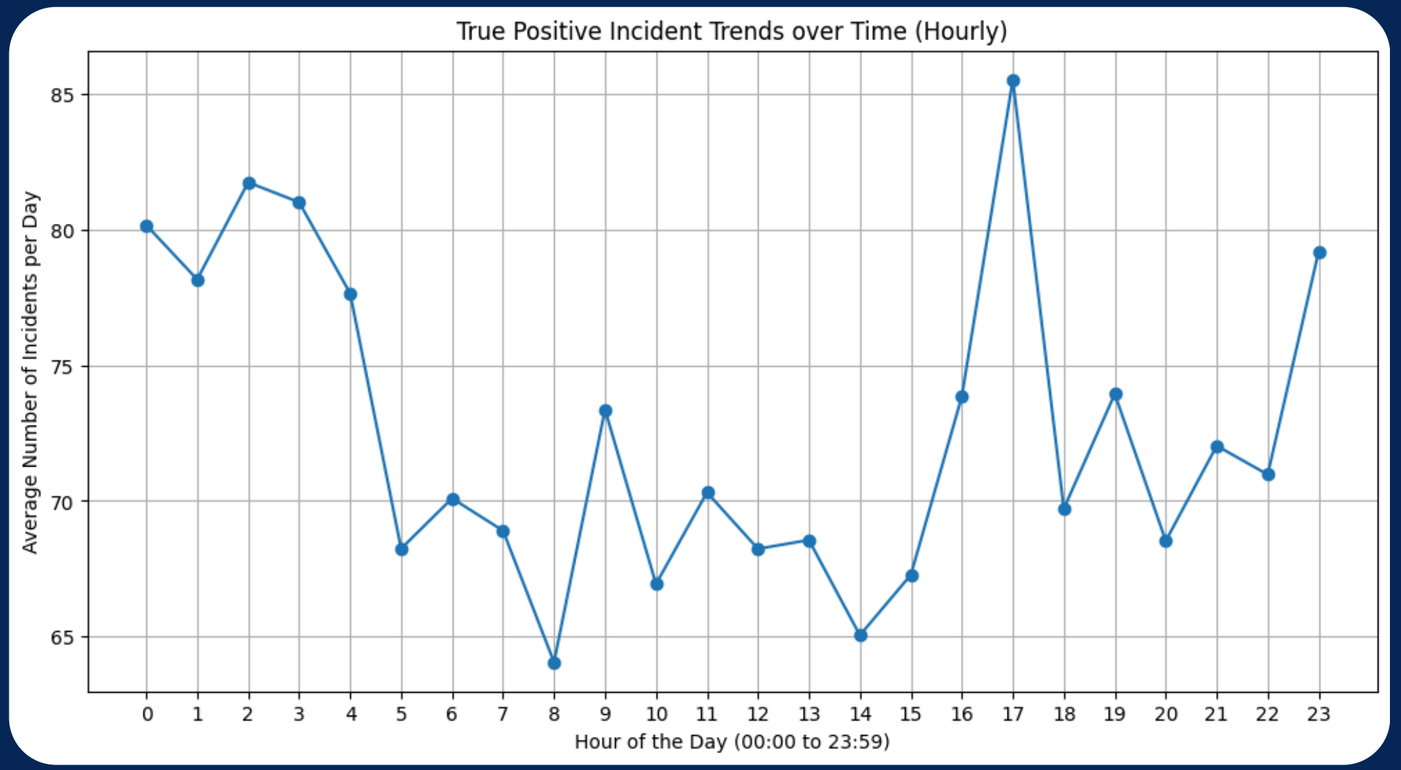


Windows



Linux

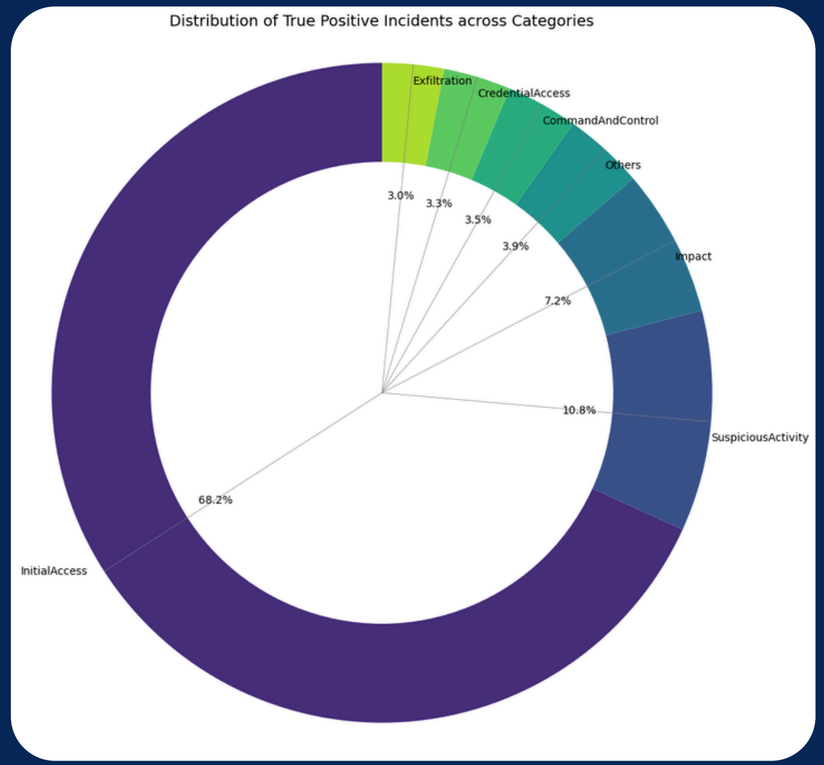
At what time do attacks occur?

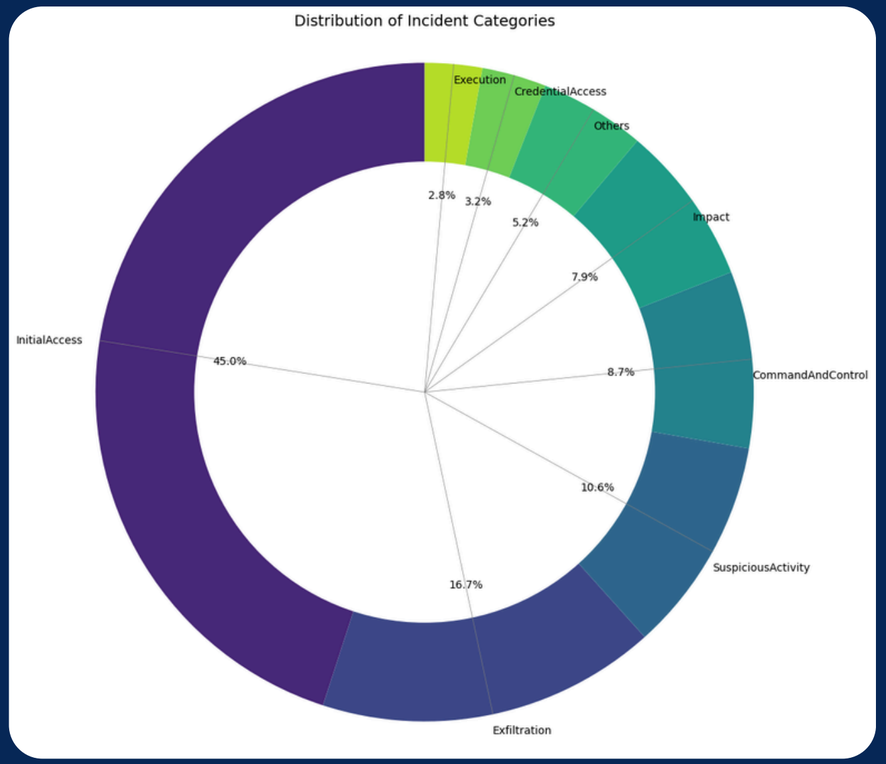


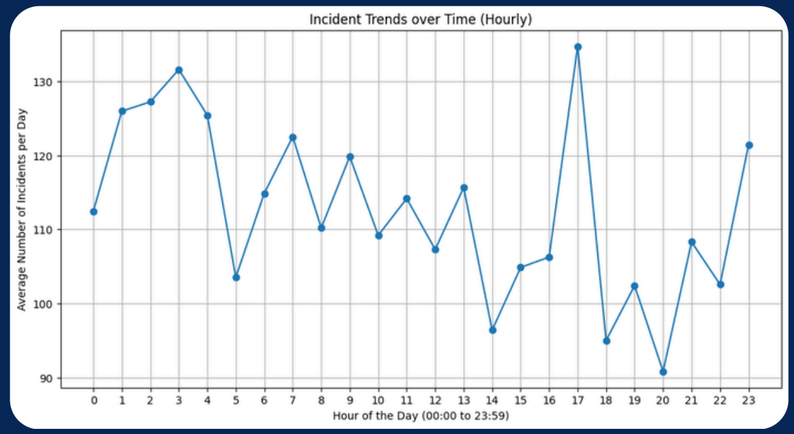
Peak 5:00 PM

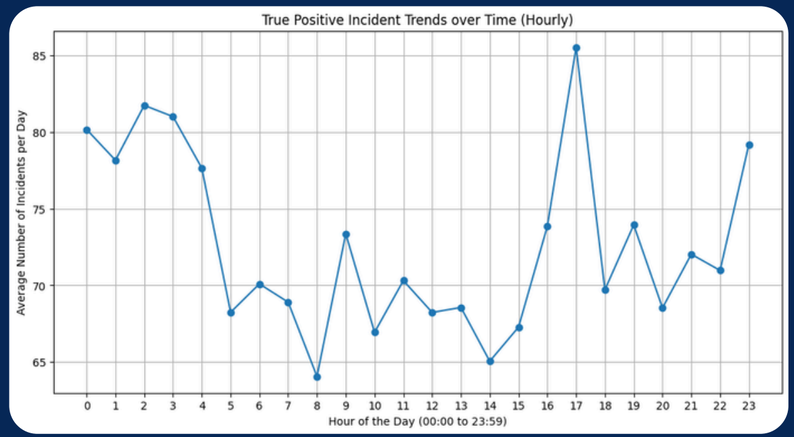
Greater during nighttime compared to daytime.

TP vs the Entire Dataset









The Need for Building a Prediction Model

Microsoft

AUTOMATION IN INCIDENT DETECTION

Trained a ML Model to Automate the Incident Detection

01. Feature Understanding & Selection

02. Model Training & Fine Tuning

03. Evaluation

Feature Selection

01. Understanding the features

02. Cleaning and Impute Missing Data

03. Select columns with a strong correlation to the Incident grade.

All columns are **categorical values.**

But few ever in form of numeric values.
Had to convert into meaningful Values

- **OS Family**:


5


1


0

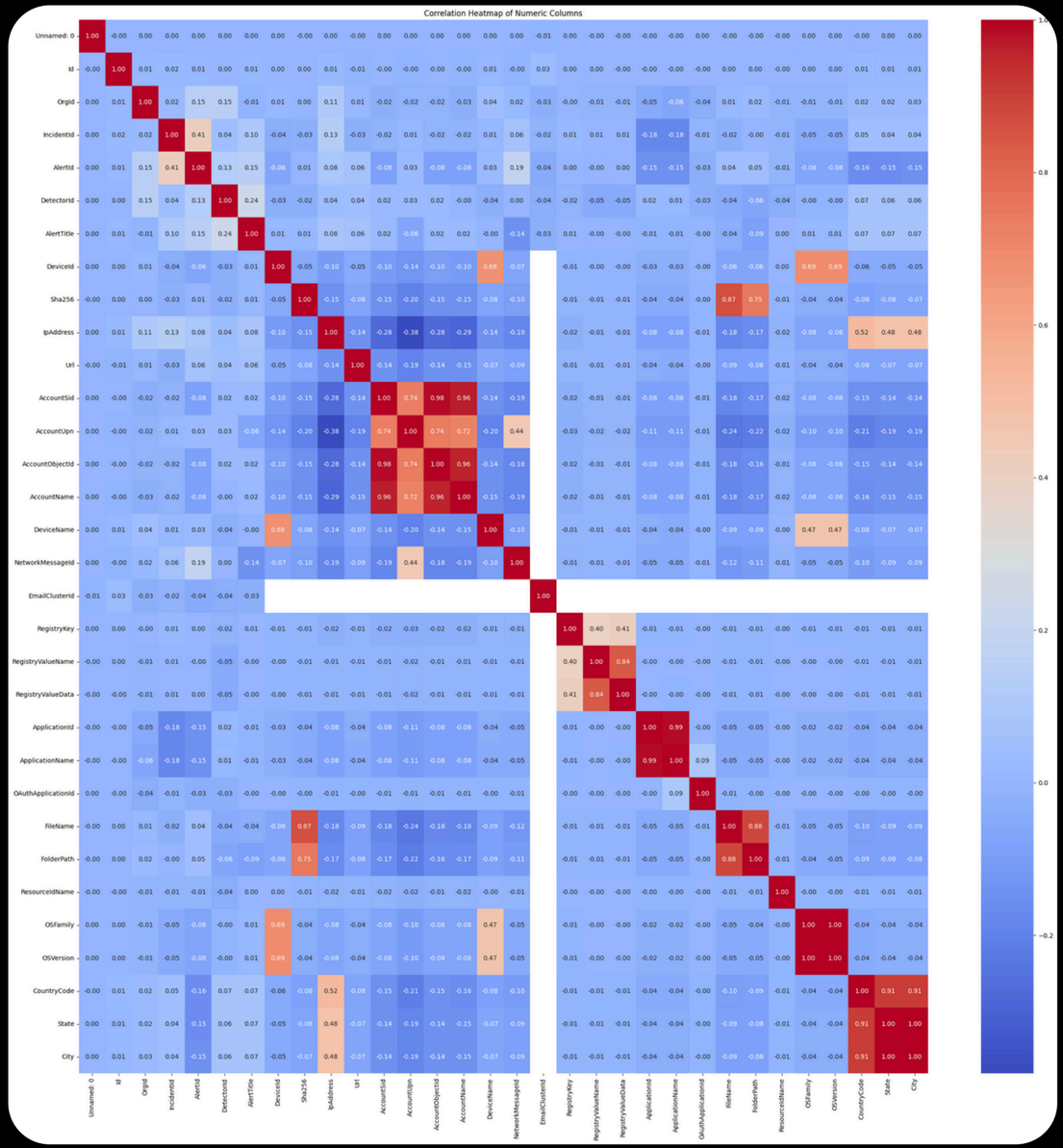
- **MitreTechniques**:
  - TA0001 - Initial Access

Feature Selection

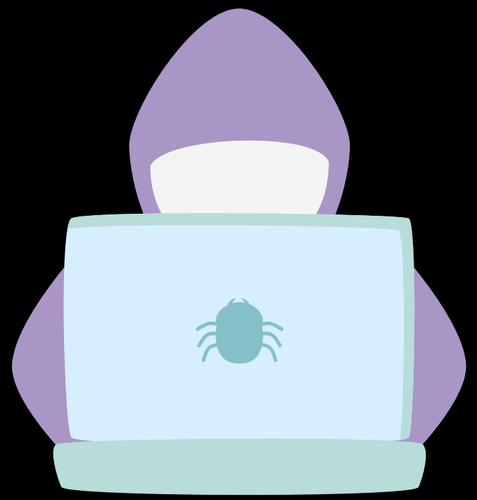
01. Understanding the features

02. Cleaning and Impute Missing Data

03. Select columns with a strong correlation to the Incident grade.



Model Architecture



Microsoft

Systems

Microsoft

Incident Detection Model

Model Training and Evaluation

Model Used for Prediction:

CATBOOST MODEL

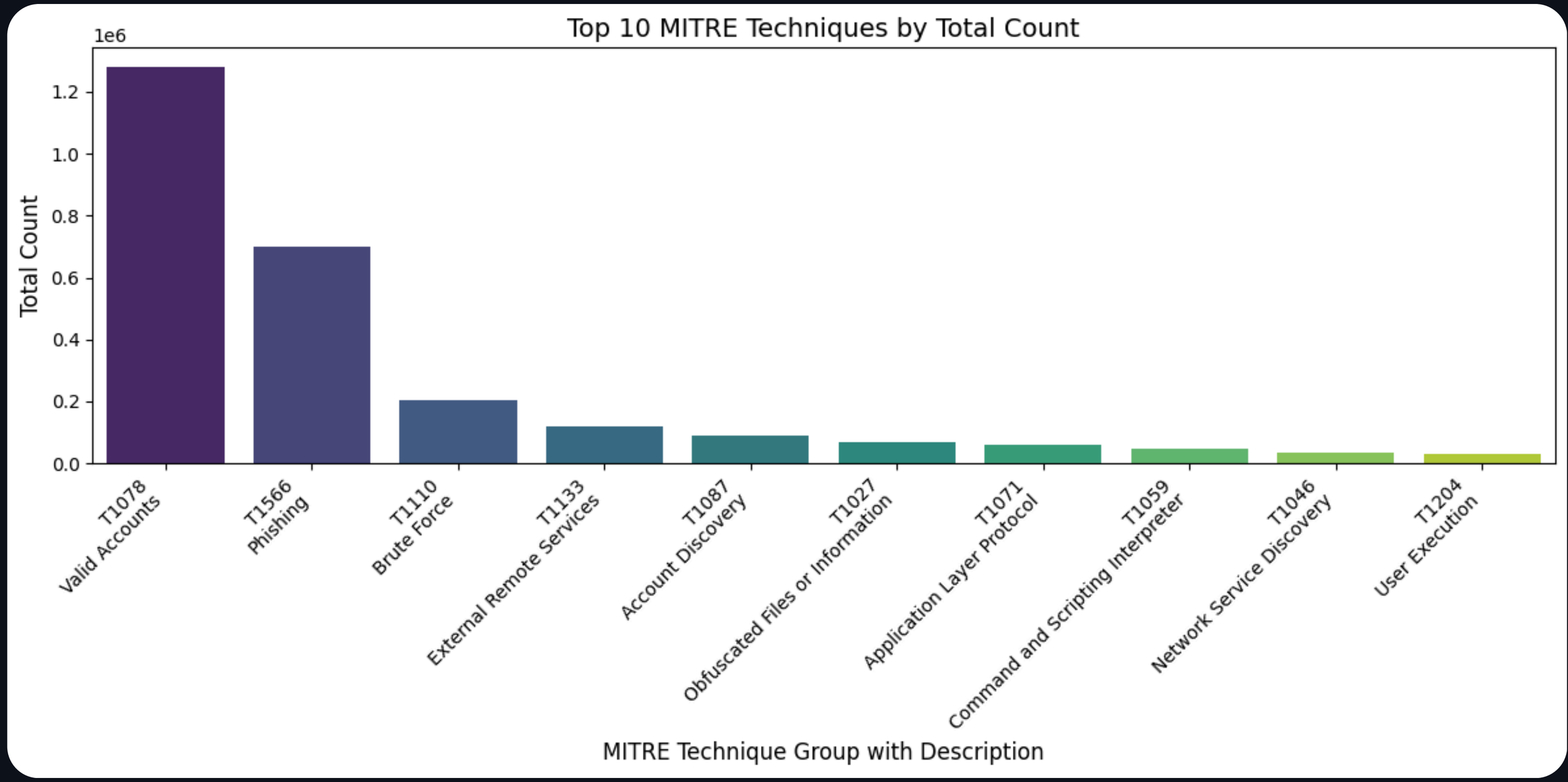
80% ACCURACY

obtained

for Incident Classification

Microsoft

TOP MITRE TECHNIQUES


Top 10 MITRE Techniques by Total Count
1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Total Count
T1078
Valid Accounts
T1566
Phishing
T1110
Brute Force
T1133
External Remote Services
T1087
Account Discovery
T1027
Obfuscated Files or Information
T1071
Application Layer Protocol
T1059
Command and Scripting Interpreter
T1046
Network Service Discovery
T1204
User Execution
MITRE Technique Group with Description

MITIGATION TECHNIQUES

- **Employee Training & Testing:** Implement yearly employee training and increase the frequency of random phishing tests
- **Least Privilege Access:** Monthly audits on server access
- **Implement Zero Trust:** Audit every admin rights access request

MITIGATION TOOLS

- **Multi-factor Authentication with Physical Access Keys:**
  - Require all employees to use hardware tokens for all assigned company computer logins
  - Require all employees to use hardware tokens or bio-metrics for all assigned phone logins
- **Supplier Software & Hardware Maintenance:** Monthly audits of supplier implemented software services & hardware for security vulnerabilities
- **Automation Incident Detection:** Implement the model presented and invest into further development

Future Directions Suggestions

- Advanced AI Integration
- Enhanced Threat Intelligence Sharing
- Zero Trust Architecture Implementation
- Proactive Cybersecurity Education
- Targeting investment to increase ROM

Return on Mitigation scoring methodology

Return on Mitigation score = (3x security value + 2x potential user impact) / Potential ease of implementation

Return on Mitigation score	Type
10 – 15	**Higher**
6 – 9	**Medium**
2 – 5	**Lower**

Percentage of users potentially impacted	Score
Lower impact (<20% of users impacted)	3
Medium impact (Up to 50% users impacted)	2
Higher impact (>50% users impacted)	1

Engagement distribution (%) by major tactic	Score
50 – 100	3
25 – 49	2
0 – 24	1

Potential ease of implementation	Score
Easy to implement (20 hours or less)	1
Medium (20 – 40 hours)	2
Harder (40+ hours)	3


Microsoft
Leading the fight
against cybercrime

THANK YOU